

123629

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Wednesday, June 02, 2004 5:44 PM
To: Holleran, Anne; STIC-Biotech/ChemLib
Subject: RE: RUSH sequence search for 09/744,197

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644 & 1642
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Holleran, Anne
Sent: Wednesday, June 02, 2004 11:27 AM
To: Chan, Christina
Subject: RUSH sequence search for 09/744,197

Please approve and forward to STIC the following RUSH sequence search request. This is for an amendment due this biweek. Thanks.

Please search the following sequences for 09/744,197:

commercial and interference databases for SEQ ID NO: 1(aa)
commercial and interference database **oligomer** search of SEQ ID NO: 1(aa)

Anne Holleran
AU: 1642
Tel: (571) 272-0833
RM: Remsen, 3A14

mailbox: Remsen, 3C18

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 3, 2004, 13:18:56 ; Search time 20 Seconds
(without alignments)
707.008 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 741
Sequence: 1 MGCMHSTODKSLHLESDPN.....LSITWMDLLQINRYKWK 147

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR.78:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	133.5	17.9	94 1 BCBOIA	S-100 protein alpha
2	131.5	17.7	94 1 BOHUIA	S-100 protein alpha
3	129.5	17.5	95 1 S35985	S-100 protein alpha
4	123	16.6	98 2 A41988	S-100 calcium-bind
5	116	15.7	97 2 A50129	S-100 protein, lun
6	106.5	14.4	101 2 A48219	calvasculin - huma
7	103.5	14.0	92 2 A48015	S-100 protein beta
8	102.5	13.8	92 2 A26557	S-100 protein beta
9	102.5	13.8	100 2 A53217	placental calcium-
10	102.5	13.8	101 2 S06207	calgrannulin B - bo
11	102	13.8	122 1 A42628	calgrannulin B - bo
12	101.5	13.7	91 1 BCBOIB	S-100 protein beta
13	99.5	13.4	101 2 S01759	calvasculin - rat
14	98.5	13.3	92 1 BCUHIB	S-100 protein beta
15	96	13.0	95 1 LUPGIB	calpactin I light
16	96	13.0	97 2 JCI139	calpactin I light
17	96	13.0	97 2 B28489	calpactin I light
18	96	13.0	113 1 UN0686	calgrannulin B - ra
19	95	12.8	97 2 A28489	calpactin I light
20	91.5	12.3	89 1 I56163	calpactin I light
21	91	12.3	97 1 JH0663	calpactin I light
22	90.5	12.2	89 1 UN0685	calgrannulin A - ra
23	89.5	12.1	95 2 S24146	S-100 protein P -
24	89	12.0	95 2 A33373	calpactin I light
25	89	12.0	114 1 B31848	calgrannulin B (val
26	86	11.6	79 1 KLBOI	calcium-binding pr
27	85.5	11.5	1162 2 T51040	hypothetical prote
28	84	11.3	79 1 KLBOI	calcium-binding pr
29	83.5	11.3	2288 2 T29999	hypothetical prote

30	83	11.2	1040	2	T19936	probable helicase
31	82.5	11.1	1390	2	T18893	hypothetical prote
32	82	11.1	407	2	T48280	hypothetical prote
33	81.5	11.0	531	2	D85059	probable calcium d
34	81.5	11.0	779	2	AE1240	translation initia
35	81.5	11.0	782	2	A11602	translation initia
36	81	10.9	113	1	S68242	calgrannulin B - mo
37	81	10.9	432	2	T36300	probable secreted
38	80.5	10.8	2493	2	S45734	probable membrane
39	80	10.8	1233	2	T40059	chromosome segrega
40	79.5	10.7	99	2	S20342	calcium-binding pr
41	79.5	10.7	174	2	F84426	hypothetical prote
42	79.5	10.7	540	1	T01989	calcium-dependent
43	79.5	10.7	571	1	F84718	hypothetical prote
44	79	10.7	110	2	B48219	S-100 calcium-bind
45	79	10.7	138	2	B81405	hypothetical prote

ALIGNMENTS

RESULT 1

BCBOIA S-100 protein alpha chain - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 14-Nov-1983 #sequence revision 06-Feb-1995 #text_change 24-Nov-1999

C/Accession: A24156; A91110; S54346; A03078

R/Kuwano, R.; Maeda, T.; Usui, H.; Araki, K.; Yamakuni, T.; Ohshima, Y.; Kurihara, T.
FEBS Lett. 202, 97-101, 1986

A/Title: Molecular cloning of cDNA of S100alpha subunit mRNA.

A/Reference number: A24156; PMID:86248083; PMID:3755105

A/Accession: A24156

A/Molecule type: mRNA

A/Residues: 1-94 <KUM>

R/Isobe, T.; Okuyama, T.

Bu. J. Biochem. 116, 79-86, 1981

A/Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a-protein

A/Reference number: A91110; PMID:8136562; PMID:7250124

A/Accession: A91110

A/Molecule type: protein

A/Residues: 2-64, 'D', 66-94 <ISO>

R/Baudier, J.; Gerard, D.

Biochemistry 22, 3360-3369, 1983

A/Title: Ions binding to S100 proteins: structural changes induced by calcium and zin

A/Reference number: A90471; PMID:8400039; PMID:6615778

A/Contents: annotation; metal ion-binding properties

R/Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.

Biochem. J. 306, 551-555, 1995

A/Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in g

A/Reference number: S54343; PMID:9519433; PMID:7887910

A/Accession: S54346

A/Molecule type: protein

A/Residues: 24-33 <OKA>

C/Comment: The S-100 protein is composed of two related polypeptide chains, alpha and

beta. The S-100 is also found in a variety of other tissues.

C/Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc

ions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

ditions, with different affinities, exist for both ions on each monomer. Physiological con

Db 3 SELETAMETLINVFHNSGKGGDKYLSKKEIKELQTELSGFLDAOKDADAVDKMKEL 62

QY 111 DEHTENKLDPEDEFMILLISTIV 132

Db 63 DENGDEVDFOEYVVALVVALTV 84

RESULT 2

BCCHUA
S-100 protein alpha chain - human
N/Alternate names: S-100 calcium-binding protein A1 (S100A1)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 23-May-1997 #text_change 22-Jun-1999
C/Accession: A44470; S16740
R/Engelkamp, D.; Schaefer, B.W.; Erne, P.; Helzlsouer, C.W.
Biochemistry 31, 10258-10264, 1992
A/Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three
A/Reference number: A44470; MUID:93041710; PMID:1384693
A/Accession: A44470
A/Molecule type: mRNA
A/Residues: 1-94 <ENG>
A/Cross-references: EMBL:X58079; NID:g36175; PID:CAA41107.1; PID:g36176
A/Experimental source: heart
A/Note: sequence extracted from NCBI backbone (NCBIP:116494)
C/Comment: This protein binds p53, tubulin and many other proteins at physiological conc
C/Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more tightly
different affinities exist for both ions on each monomer. Physiological concentrations
binding sites.
C/Comment: Although predominant among the water-soluble brain proteins, S-100 is also found
C/Genetics: GDB:S100A1; S100A
A/Cross-references: GDB:126839; OMIM:176940
A/Map position: 1q21-1q21
C/Complex: homodimer; heterodimer with S-100 protein beta chain (see PIR:BCCHUB)
C/Superfamily: S-100 protein; calmodulin repeat homology
C/Keywords: brain; calcium binding; EF hand; heterodimer; homodimer; zinc
F/-41/Domain: calmodulin repeat homology <EF1>
F/50-82/Domain: calmodulin repeat homology <EF2>

Query Match 17.7%; Score 131.5; DB 1; Length 94;
Best Local Similarity 32.9%; Pred. No. 0.00017;
Matches 27; Conservative 23; Mismatches 29; Indels 3; Gaps 1;

QY 54 SDLEKAIATTAIFRNSSDSG---KLEKAIKDLIQGFNFPAEGQETKPKYREILSEL 110

Db 3 SELETAMETLINVFHNSGKGGDKYLSKKEIKELQTELSGFLDAOKDADAVDKMKEL 62

QY 111 DEHTENKLDPEDEFMILLISTIV 132

Db 63 DENGDEVDFOEYVVALVVALTV 84

RESULT 3

S35985
S-100 protein alpha chain - weatherfish
C/Species: Misgurnus fossilis (weatherfish)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: S35985
R/Ivanenko, V.V.; Gerke, V.; Minin, A.A.; Plessmann, U.; Weber, K.
Mech. Dev. 42, 151-158, 1993
A/Title: Transduction of Ca(2+) signals upon fertilization of eggs; identification of an
A/Reference number: S35985; MUID:94031845; PMID:8217841
A/Accession: S35985
A/Molecule type: protein
A/Residues: 1-95 <IVA>
C/Superfamily: S-100 protein; calmodulin repeat homology
C/Keywords: calcium binding; EF hand
F/49-81/Domain: calmodulin repeat homology <EF2>

Query Match 17.5%; Score 129.5; DB 1; Length 95;
Best Local Similarity 29.8%; Pred. No. 0.00026;
Matches 28; Conservative 27; Mismatches 34; Indels 5; Gaps 2;

QY 54 SDLEKAIATTAIFRNSSDSG---KLEKAIKDLIQGFNFPAEGQETKPKYREILSEL 110

Db 2 SLESAMESLIVFHTYSSKEGDKYLSKKEIKELQTELSGFLDAOKDADAVDKMKEL 61

QY 111 DEHTENKLDPEDEFMILLISTIVMSD--LLQINRN 142

Db 62 DENGDEVDFOEYVVALVVALTVACNEFFLESKN 95

RESULT 4

A41988
S-100 calcium-binding protein A2 - human
N/Alternate names: calcium-binding protein CAN19; S-100 calcium-binding protein L
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Aug-1997
C/Accession: A41988
R/Lea, S.W.; Tomasetto, C.; Wisshelm, K.; Keyomarsi, K.; Sager, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 2504-2508, 1992
A/Title: Down-regulation of a member of the S100 gene family in mammary carcinoma cell
A/Reference number: A41988; MUID:92196147; PMID:1372446
A/Accession: A41988
A/Molecule type: mRNA
A/Residues: 1-98 <LEB>
A/Cross-references: GB:M87068; NID:q179896
A/Experimental source: mammary carcinoma cells
A/Note: sequence extracted from NCBI backbone (NCBIN:88561, NCBIP:88562)
C/Genetics: GDB:S100A2; CAN19; S100L
A/Cross-references: GDB:211118; OMIM:176993
A/Map position: 1q21-1q21
C/Superfamily: S-100 protein; calmodulin repeat homology
C/Keywords: calcium binding; EF hand
F/8-42/Domain: calmodulin repeat homology <EF1>
F/51-83/Domain: calmodulin repeat homology <EF2>

Query Match 16.6%; Score 123; DB 2; Length 98;
Best Local Similarity 31.1%; Pred. No. 0.00097;
Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;

QY 53 CSDEKAIATTAIFRNSSDSG---KLEKAIKDLIQGFNFPAEGQETKPKYREILSEL 109

Db 3 CSSEQLALVAVTTFKYSQSGDKPKLSKKEIKELHLPFVGEKVDSEGLKLMGN 62

QY 110 DEHTENKLDPEDEFMILLISTIV--SDLLQ 138

Db 63 LDENSQQVDFOEYVVALVVALTVGNDFFQ 92

RESULT 5

A30129
S-100 protein, lung - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Sep-1995
C/Accession: A30129
R/Glenney Jr., J.R.; Kindy, M.S.; Zokas, L.
J. Cell Biol. 108, 569-578, 1989
A/Title: Isolation of a new member of the S100 protein family: amino acid sequence,
A/Reference number: A30129; MUID:89139574; PMID:2521861
A/Accession: A30129
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-97 <GLB>
C/Superfamily: S-100 protein; calmodulin repeat homology
C/Keywords: Calcium binding; EF hand
F/50-82/Domain: calmodulin repeat homology <EF2>

Query Match 15.7%; Score 116; DB 2; Length 97;
Best Local Similarity 29.2%; Pred. No. 0.0038;
Matches 26; Conservative 24; Mismatches 35; Indels 4; Gaps 2;

QY 54 SDLEKAIATTAIFRNSSDSG---KLEKAIKDLIQGFNFPAEGQETKPKYREILSEL 110

Db 3 SPLDQALVAVTTFKYSQSGDKPKLSKKEIKELHLPFVGEKVDSEGLKLMGN 62

C:Keywords: blocked amino end; calcium binding; EF hand; heterodimer; inflammation; phosph
F;6-40/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 13.8%; Score 102; DB 1; Length 122;
Best Local Similarity 31.8%; Pred. No. 0.078;
Matches 27; Conservative 19; Mismatches 35; Indels 4; Gaps 2;

QY 54 SDLEKAIATTAIFRNSSDGKLEKAI--AKDLQTFRNFAGQ-ETKPKREILSE 109
DB 2 SQMSESEITINIFHQYVRLGHVDTLQKESQVQKELPFLKQKKEALNEMED 61
QY 110 LDHTENKLDPEDEPMLLSITVMS 134
DB 62 LDTNVKQSLSEFFIMLVARLVAS 86

RESULT 12

BCH01B

S-100 protein beta chain - bovine
N:Alternate names: neurocalcin delta-binding protein S100-beta
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-May-1979 #sequence, revision 14-Nov-1983 #text_change 24-Nov-1999
C:Accession: A91254; B91110; A90075; S54348; A03077
R:Isoe, T.; Okuyama, T.
Eur. J. Biochem. 89, 379-388, 1978
A>Title: The amino-acid sequence of S-100 protein (PAP-I-b protein) and its relation to
A:Reference number: A91254; PMID:79045285; PMID:710395
A:Accession: A91254
A:Molecule type: protein
A:Residues: 'ESEL', 5-91 <ISO>
A:Experimental source: brain
A>Note: This sequence has since been revised in reference A91110
R:Isoe, T.; Okuyama, T.
Eur. J. Biochem. 116, 79-86, 1981
A>Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.
A:Reference number: A91110; PMID:81236562; PMID:7250124
A:Accession: B91110
A:Molecule type: protein
A:Residues: 1-91 <ISO>
R:Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A>Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc c
A:Reference number: A90471; PMID:8400039; PMID:6615778
A:Contents: annotation; metal ion-binding properties
R:Marshall, D.R.; Umekawa, H.; Waterson, D.W.; Hidaka, H.
Arch. Biochem. Biophys. 240, 777-780, 1985
A>Title: Structural characterization of the calcium binding protein S100 from adipose ti
A:Reference number: A90075; PMID:85278169; PMID:4026904
A:Accession: A90075
A:Molecule type: protein
A:Residues: 1-91 <MAR>
A:Experimental source: adipose tissue
R:Okazaki, K.; Obara, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A>Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A:Reference number: S54343; PMID:9514333; PMID:7887910
A:Accession: S54348
A:Molecule type: protein
A:Residues: 35-61, 'N', 63-79, 'V' <OKA>
C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and be
C:Comment: S-100 is also found in a variety of other tissues.
C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve
es with different affinities exist for both ions on each monomer. Physiological concentr
cium-binding sites.
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
F;5-39/Domain: calmodulin repeat homology <EF1>
F;48-80/Domain: calmodulin repeat homology <EF2>
F;1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental
F;18,21,23,26,31/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F;61,63,65,67,72/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted

Query Match 13.7%; Score 101.5; DB 1; Length 91;
Best Local Similarity 28.4%; Pred. No. 0.061;
Matches 23; Conservative 20; Mismatches 35; Indels 3; Gaps 1;

QY 54 SDLEKAIATTAIFRNSSDG--KLEKAIKDLQTFRNFAGQETKPKREILSEL 110
DB 1 SELEKVALIVFQYSGREDDKRLKSEKLEKLEINLSHFLKEIKQEVVDKVVETL 60
QY 111 DEHTENKLDPEDEPMLLSIT 131
DB 61 DSDGDECDPFRNFVAMIT 81

RESULT 13

S01759

calvasculin - rat
N:Alternate names: calcium-binding protein p9Ka/42A; gene p9Ka protein
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1989 #sequence, revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: S01759; B31373; S26496; S26497
R:Barclough, R.; Savin, U.; Dube, S.K.; Rudland, P.S.
J. Mol. Biol. 198, 13-20, 1987
A>Title: Molecular cloning and sequence of the gene for p9Ka : a cultured myocytethe1
A:Reference number: S01759; PMID:88118907; PMID:3430604
A:Accession: S01759
A:Molecule type: DNA
A:Residues: 1-101 <BAR>
A:Cross-references: EMBL:X06916; NID:956832; PIDD:CAA30014.1; PID:956833
R:Maslakowski, P.; Shooter, E.M.
Proc. Natl. Acad. Sci. U.S.A. 85, 1277-1281, 1988
A>Title: Nerve growth factor induces the genes for two proteins related to a family c
A:Reference number: A94189; PMID:88125019; PMID:3422491
A:Accession: B31373
A:Molecule type: mRNA
A:Residues: 1-101 <MAS>
A:Cross-references: GB:U03628; NID:9206829; PIDD:AAA42098.1; PID:9206830
R:de Voege, N.W.; Kerkhjee, B.B.
Oncogene 7, 109-119, 1992
A>Title: Transformation of normal rat kidney cells by v-K-ras enhances expression of
A:Reference number: S26496; PMID:92158347; PMID:1741158
A:Accession: S26496
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 65-100 <DEV>
A:Cross-references: EMBL:X64023
A:Accession: S26497
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 3-97 <DE2>
A:Cross-references: EMBL:X64022
C:Genetics:
A:Gene: p9Ka
A:Insertions: 47/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 13.4%; Score 99.5; DB 2; Length 101;
Best Local Similarity 25.9%; Pred. No. 0.1;
Matches 21; Conservative 24; Mismatches 33; Indels 3; Gaps 1;

QY 56 LEKAIATTAIFRNSSDG--KLEKAIKDLQTFRNFAGQETKPKREILSEL 112
DB 5 LBEALDLVISTHRSXSGEGRFKINKTELKELTRLEPFLGRTOBAAPQKMMNDS 64
QY 113 HTENKLDPEDEPMLLSITVMS 133
DB 65 NRDNEDVDFQYCVFLSCIAMM 85

RESULT 14

BCH01B

S-100 protein beta chain (validated) - human

N/Alternate names: neural S-100 calcium-binding protein beta
 C/Species: Homo sapiens (man)
 C/Date: 04-Dec-1986 #sequence revision 06-Jan-1995 #text_change 08-Dec-2000
 C/Accession: A28364; A2972; A03076
 R/Allorg: R.J.; Friend, W.C.; O'Hallion, D.; Neilson, K.M.; Baumal, R.; Dunn, R.J.; Marke
 J. Biol. Chem. 265, 15537-15543, 1990
 A/Title: Cloning and expression of the human S100beta gene.
 A/Reference number: A28364; PMID:90368757; PMID:2394738
 A/Accession: A28364
 A/Molecule type: DNA
 A/Residues: 1-92 <ALT>
 A/Cross-references: GB:J05600; GB:M59486; NID:9337726; GB:M59487; NID:9337727; GB:M59488
 R/Jensen, R.; Marshak, D.R.; Anderson, C.; Lukac, T.J.; Watterson, D.W.
 J. Neurochem. 45, 700-705, 1985
 A/Title: Characterization of human brain S100 protein fraction: amino acid sequence of S
 A/Reference number: A2972; PMID:85291729; PMID:4031854
 A/Accession: A2972
 A/Molecule type: protein
 A/Residues: 2-92 <GEN>
 R/Bader, J.; Glaeser, N.; Haglid, K.; Gerard, D.
 Biochim. Biophys. Acta 790, 164-173, 1984
 A/Title: Purification, characterization and ion binding properties of human brain S100b
 A/Reference number: A20653; PMID:85023393; PMID:6487634
 A/Accession: A20653
 C/Contents: annotation; metal ion-binding properties
 C/Comment: This protein binds p53, tubulin and many other proteins at physiological cond
 C/Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more tigh
 different affinities exist for both ions on each monomer. Physiological concentrations
 nding sites.
 C/Comment: This protein is expressed predominantly in brain tissue by astroglial cells.
 C/Comment: The homodimer contains disulfide bonds, but the bond pattern has not been det
 C/Genetics:
 A/Genes: GDB:S100B
 A/Cross-references: GDB:120360; OMIM:176990
 A/Map position: 21q22.3-21q22.3
 A/Introns: 46/3
 A/Note: the first intron occurs before the initiator codon
 C/Complex: homodimer; heterodimer with S-100 protein alpha chain (see PIR:BCHTUA)
 C/Superfamily: S-100 protein; calmodulin repeat homology
 C/Keywords: blocked amino end; brain; calcium binding; EF hand; heterodimer; homodimer;
 F/2-92/Product: S-100 protein beta chain #status experimental <MAT>
 F/6-40/Domain: calmodulin repeat homology <EF1>
 F/49-81/Domain: calmodulin repeat homology <EF2>
 F/2/Modified sites: blocked amino end (Ser) (in mature form) (probably acetylated) #statu
 F/19/22/24/27/32/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
 F/62/64/66/73/Binding site: calcium (Asp, Asp, Asp, Glu, Glu) #status predicted

Query Match 13.3%; Score 98.5; DB 1; Length 92;
 Best Local Similarity 27.2%; Pred. No. 0.11;
 Matches 22; Conservative 21; Mismatches 35; Indels 3; Gaps 1;

QY 54 SLEKAIATLALIFRNSDSDG---KLEKAIKDLQTOFRNFAEGQETKPKYREILSEL 110
 Db 2 SLEKAIATLALIDVFQYSGREGDKKLEKSELKELINNELSHFLERIKEDQVVDKWEEL 61

QY 111 DEHTENKLDPEDDPMILLSTIV 131
 Db 62 DNDGDGDCDFCEFMFAFVAMVT 82

RESULT 15
 LURP10
 calpactin I light chain - pig
 N/Alternate names: p10 protein; p11 protein
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 18-Jul-1997
 C/Accession: A03079; A24063
 R/Gerke, V.; Weber, K.
 EMBO J. 4, 2917-2920, 1985
 A/Title: The regulatory chain in the p16-kD substrate complex of viral tyrosine-specific
 A/Reference number: A03079; PMID:86053730; PMID:2996764
 A/Accession: A03079
 A/Molecule type: protein
 A/Residues: 1-95 <GER>

R/Hexham, J.M.; Totty, N.F.; Waterfield, M.D.; Crumpton, M.J.
 Biochem. Biophys. Res. Commun. 134, 248-254, 1986
 A/Title: Homology between the subunits of S100 and A 10kDa polypeptide associated wit
 A/Reference number: A24063; PMID:86130472; PMID:2936341
 A/Accession: A24063
 A/Molecule type: protein
 A/Residues: 1-51, 'XX', 54-55 <HEX>
 C/Comment: Calpactin I light does not appear to bind calcium.
 C/Comment: Calpactin I is a tetramer of two light chains and two heavy chains (annex)
 C/Function:
 A/Description: found in lamina beneath plasma membrane where it may cross-link plasm
 A/Pathway: exocytosis
 C/Superfamily: S-100 protein; calmodulin repeat homology
 C/Keywords: EF hand; intestine
 F/6-37/Domain: calmodulin repeat homology <EF1>
 F/46-78/Domain: calmodulin repeat homology <EF2>

Query Match 13.0%; Score 96; DB 1; Length 95;
 Best Local Similarity 26.6%; Pred. No. 0.19;
 Matches 21; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

QY 54 SLEKAIATLALIFRNSDSDGKLEKAIKDLQTOFRNFAEGQETKPKYREILSEL 113
 Db 2 SLEKAIATLALIDVFQYSGREGDKKLEKSELKELINNELSHFLERIKEDQVVDKWEEL 61

QY 114 DEHTENKLDPEDDPMILLSTIV 132
 Db 62 DNDGDGDCDFCEFMFAFVAMVT 80

Search completed: June 3, 2004, 13:22:27
 Job time: 21 secs

Fri Jun 4 16:16:00 2004

US-09-744-197-1.01ligo.rat

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:29:03 ; Search time 23 Seconds
(without alignments)
329.957 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 147
Sequence: 1 MGGCMHSTQDKSLHLESDPN.....LSITWMDLLQINIRVKIMK 147

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size: 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	147	US-09-621-976-7168	Sequence 7168, Ap
2	8	5.4	163	US-09-252-991A-24968	Sequence 24968, A
3	8	5.4	361	US-09-543-681A-5180	Sequence 5180, Ap
4	7	4.8	126	US-09-134-001C-3766	Sequence 3766, Ap
5	7	4.8	130	US-09-252-991A-21537	Sequence 21537, A
6	7	4.8	232	US-09-252-991A-26162	Sequence 26162, A
7	7	4.8	262	US-09-107-532A-5791	Sequence 5791, Ap
8	7	4.8	385	US-09-489-039A-13370	Sequence 13370, A
9	7	4.8	479	US-08-416-788-2	Sequence 2, Appl
10	7	4.8	816	US-08-190-802A-54	Sequence 54, Appl
11	7	4.8	816	US-08-477-346-54	Sequence 54, Appl
12	7	4.8	816	US-08-473-089-54	Sequence 54, Appl
13	7	4.8	816	US-08-487-072A-54	Sequence 54, Appl
14	7	4.8	1311	US-08-540-406-4	Sequence 4, Appl
15	7	4.8	1311	US-08-556-055-4	Sequence 4, Appl
16	7	4.8	1311	US-08-954-668-4	Sequence 4, Appl
17	7	4.8	1311	US-08-918-658-4	Sequence 4, Appl
18	7	4.8	1311	US-08-724-631-4	Sequence 4, Appl
19	7	4.8	1311	US-08-954-701A-4	Sequence 4, Appl
20	7	4.8	1311	PCT-US95-1333-4	Sequence 4, Appl
21	6	4.1	26	US-08-346-147B-45	Sequence 45, Appl
22	6	4.1	26	US-08-497-214D-45	Sequence 45, Appl
23	6	4.1	27	US-09-480-993-2	Sequence 2, Appl
24	6	4.1	27	US-09-367-065A-2	Sequence 2, Appl
25	6	4.1	27	US-09-568-466-2	Sequence 2, Appl
26	6	4.1	34	US-09-447-158B-23	Sequence 23, Appl
27	6	4.1	40	US-08-188-228-12	Sequence 12, Appl

28	6	4.1	40	1	US-08-332-643-12	Sequence 12, Appl
29	6	4.1	40	1	US-08-332-638-12	Sequence 12, Appl
30	6	4.1	40	4	US-09-480-993-18	Sequence 18, Appl
31	6	4.1	67	4	US-09-543-681A-4451	Sequence 4451, Ap
32	6	4.1	72	4	US-09-673-395A-527	Sequence 527, App
33	6	4.1	73	4	US-09-621-976-3908	Sequence 3908, Ap
34	6	4.1	83	2	US-08-465-380-52	Sequence 52, Appl
35	6	4.1	83	2	US-08-486-397-52	Sequence 52, Appl
36	6	4.1	83	2	US-08-486-399-52	Sequence 52, Appl
37	6	4.1	83	2	US-08-461-965-52	Sequence 52, Appl
38	6	4.1	83	2	US-08-634-641-52	Sequence 52, Appl
39	6	4.1	83	3	US-09-249-471-52	Sequence 52, Appl
40	6	4.1	83	3	US-09-249-472-52	Sequence 52, Appl
41	6	4.1	83	3	US-09-249-451-52	Sequence 52, Appl
42	6	4.1	83	3	US-08-809-455-52	Sequence 52, Appl
43	6	4.1	83	3	US-09-249-461-52	Sequence 52, Appl
44	6	4.1	83	3	US-09-249-448-52	Sequence 52, Appl
45	6	4.1	83	4	US-09-249-473-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-09-621-976-7168
; Sequence 7168, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT FILING DATE: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO: 7168
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7168

Query Match 100.0%; Score 147; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.5e-143; Indels 0; Gaps 0;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGCMHSTQDKSLHLESDPNPSAAPTSTCAERKPKRISISKOLASVAKLAKGSDLEKAI 60
DB 1 MGGCMHSTQDKSLHLESDPNPSAAPTSTCAERKPKRISISKOLASVAKLAKGSDLEKAI 60
QY 61 ATTALIFNSDSDGKLEKATKADLOTFQFNPAEGGRTKRYREIISLDEHTENKIDF 120
DB 61 ATTALIFNSDSDGKLEKATKADLOTFQFNPAEGGRTKRYREIISLDEHTENKIDF 120
QY 121 EDFMILLSTWMDLLQINIRVKIMK 147
DB 121 EDFMILLSTWMDLLQINIRVKIMK 147
RESULT 2
US-09-252-991A-24968
; Sequence 24968, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 10/196.116
; CURRENT FILING DATE: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

Fri Jun 4 16:16:00 2004

us-09-744-197-1.01igo.ra1

Page 2

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24968
LENGTH: 163
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24968

Query Match
Best Local Similarity 5.4%; Score 8; DB 4; Length 163;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 PSAAPTS 28
DB 52 PSAAPTS 59

RESULT 3
US-09-543-681A-5180
Sequence 5180, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709,1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5180
LENGTH: 361
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5180

Query Match
Best Local Similarity 5.4%; Score 8; DB 4; Length 361;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 MILLIST 131
DB 167 MILLIST 174

RESULT 4
US-09-134-001C-3766
Sequence 3766, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3766
LENGTH: 126
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3766

Query Match
Best Local Similarity 4.8%; Score 7; DB 4; Length 126;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 ISKOLAS 46
DB 16 ISKOLAS 22

RESULT 5
US-09-252-991A-21537
Sequence 21537, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21537
LENGTH: 130
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21537

Query Match
Best Local Similarity 4.8%; Score 7; DB 4; Length 130;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 PSAAPTS 27
DB 65 PSAAPTS 71

RESULT 6
US-09-252-991A-26162
Sequence 26162, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26162
LENGTH: 232
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26162

Query Match
Best Local Similarity 4.8%; Score 7; DB 4; Length 232;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 PSAAPTS 27
DB 65 PSAAPTS 71

RESULT 7
US-09-107-532A-5791
Sequence 5791, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush

Fri Jun 4 16:16:00 2004

us-09-744-197-1.01igo.rat

Page 3

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arindello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5791:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...262
SEQUENCE DESCRIPTION: SEQ ID NO: 5791:
US-09-107-532A-5791
Query Match 4.8%; Score 7; DB 4; Length 262;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;
QY 125 ILLSIT 131
DB 21 ILLSIT 27
RESULT 8
US-09-489-039A-13370
Sequence 13370, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13370
LENGTH: 385
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13370

Query Match 4.8%; Score 7; DB 4; Length 385;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;
QY 43 QLASVKA 49
DB 217 QLASVKA 223
RESULT 9
US-08-416-788-2
Sequence 2, Application US/08416788
Patent No. 5780245
GENERAL INFORMATION:
APPLICANT: Marteau, Luc
TITLE OF INVENTION: No. 5780245el Polypeptides Having a Serotonin
RECEPTOR ACTIVITY, Nucleic Acids Coding for These
TITLE OF INVENTION: Receptor Activity, Nucleic Acids Coding for These
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,788
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/01012
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92-12280
FILING DATE: 14-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K. 619
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: EX92008-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-788-2
Query Match 4.8%; Score 7; DB 1; Length 479;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;
QY 46 SVKALRK 52
DB 403 SVKALRK 409
RESULT 10
US-08-190-802A-54
Sequence 54, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
US-08-190-802A-54

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: MUS MUSCULUS PROTEIN, Fig. 37
US-08-190-802A-54
Query Match 4.8%; Score 7; DB 1; Length 816;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 NSSSDG 75
DB 644 NSSSDG 650
RESULT 11
US-08-477-346-54
Sequence 54, Application US/08477346
Patent No. 6262023
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514

PRIOR APPLICATION DATA: 08/487,072
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: MUS MUSCULUS PROTEIN, Fig. 37
US-08-477-346-54
Query Match 4.8%; Score 7; DB 3; Length 816;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 NSSSDG 75
DB 644 NSSSDG 650
RESULT 12
US-08-473-089-54
Sequence 54, Application US/08473089
Patent No. 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,089
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: MUS MUSCULUS PROTEIN, Fig. 37
US-08-473-089-54

Query Match 4.8%; Score 7; DB 4; Length 816;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 NSSSDSG 75
DB 644 NSSSDSG 650

RESULT 13
US-08-487-072A-54
Sequence 54, Application US/08487072A
Patent No. 6423684
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thecof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: MUS MUSCULUS PROTEIN, Fig. 37
US-08-487-072A-54
Query Match 4.8%; Score 7; DB 4; Length 816;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 NSSSDSG 75
DB 644 NSSSDSG 650
RESULT 14
US-08-540-406-4
Sequence 4, Application US/08540406
Patent No. 5837538
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V

APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540,406
FILING DATE: 06-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-540-406-4

Query Match 4.8%; Score 7; DB 2; Length 1311;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 LLISITV 132
DB 468 LLISITV 474

RESULT 15
US-08-656-055-4
Sequence 4, Application US/08656055
Patent No. 6027882
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/540,406
FILING DATE:

Fri Jun 4 16:16:00 2004

us-09-744-197-1.oligo.rat

Page 6

ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Berttram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-656-055-4

Query Match 4.8%; Score 7; DB 3; Length 1311;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 LLLSTTV 132
Db 468 LLLSTTV 474

Search completed: June 3, 2004, 13:34:46
Job time: 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:22:32 ; Search time 57 Seconds
(without alignments)

728.675 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 147
Sequence: 1 MGACGHRSTQDKSLHLEBDN.....LSITVMSDLLQINIKIMK 147

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

1: Genesegp296a04:*
2: Genesegp19808:*
3: Genesegp19908:*
4: Genesegp20008:*
5: Genesegp20018:*
6: Genesegp20038:*
7: Genesegp20038:*
8: Genesegp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	147	3	AA58793 Human cal
2	143	97.3	143	4	AA40026 Human pol
3	138	93.9	140	4	AA115494 Human nov
4	138	93.9	140	6	ABUS5563 Human nov
5	137	93.2	137	4	AA41812 Human pol
6	8	5.4	161	3	AA333229 Eucalyptu
7	8	5.4	229	2	AA99637 Eucalyptu
8	8	5.4	314	5	AA25918 Pokeweed
9	8	5.4	572	3	AA32036 Arabidops
10	8	5.4	1577	4	AB59387 Drosophil
11	8	4.8	41	4	AB42868 Peptide #
12	8	4.8	41	4	AA36682 Peptide #
13	8	4.8	41	4	AB26140 Protein #
14	8	4.8	41	4	AA476575 Human bon
15	8	4.8	41	4	AA63761 Human bon
16	8	4.8	41	4	ABG58274 Human liv
17	8	4.8	41	5	ABG45823 Human pep
18	8	4.8	56	6	AA43681 Propionib
19	8	4.8	56	6	AB440200 Propionib
20	8	4.8	126	5	AB38921 Staphyloc
21	8	4.8	127	4	AA95510 Human rep
22	8	4.8	127	4	AB36193 Human tes
23	8	4.8	133	5	AB91269 Herbicida
24	8	4.8	138	5	ABP04851 Human ORF
25	8	4.8	145	5	ABP69561 Human pol

26	7	4.8	160	5	ABP03296 Human ORF
27	7	4.8	162	6	AA55057 Human pro
28	7	4.8	179	4	AB62818 Drosophil
29	7	4.8	188	5	AB74513 Human col
30	7	4.8	198	5	AB93224 C. albica
31	7	4.8	207	6	AA49163 Propionib
32	7	4.8	213	3	AA45682 Propionib
33	7	4.8	257	3	AB16291 Pinus rad
34	7	4.8	260	3	AA33135 Zea mays
35	7	4.8	262	7	AA33134 Zea mays
36	7	4.8	282	6	AD09464 E. faeciu
37	7	4.8	282	6	AB06746 Altiococ
38	7	4.8	283	3	AA33133 Zea mays
39	7	4.8	285	3	AA33133 Arabidops
40	7	4.8	343	3	AA33133 Arabidops
41	7	4.8	346	3	AA33133 Arabidops
42	7	4.8	380	3	AA33133 Arabidops
43	7	4.8	380	3	AA33133 Arabidops
44	7	4.8	381	6	AB06748 Altiococ
45	7	4.8	382	3	AA33133 Arabidops

ALIGNMENTS

RESULT 1	AA58793	standard; protein, 147 AA.
ID	AA58793	
XX	AA58793	
AC	AA58793	
XX	08-MAY-2000	(first entry)
DT		
XX		
DE	Human calcium regulatory protein CaREG-1.	
XX		
KW	Human; calcium regulatory protein 1; CaREG-1; anticonvulsant; vasotrophic;	
KW	cerebroprotective; neurotrophic; neuroprotective; neuroleptic; cardiant;	
KW	cytostatic; epilepsy; stroke; Alzheimer's disease; cardiomyopathy;	
KW	myocarditis; Duchenne's muscular dystrophy; tachyarrhythmia; cancer;	
KW	diagnosis; therapy.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Region	1..98
FT	Modified-site	/note= "similarity to annexin type I signature"
FT	Modified-site	7
FT	Modified-site	/note= "potential O-phosphorylation"
FT	Modified-site	39
FT	Modified-site	/note= "potential O-phosphorylation"
FT	Modified-site	54
FT	Modified-site	/note= "potential O-phosphorylation"
FT	Region	/note= ".144
FT	Region	/note= "similarity to S100 Ca2+ binding protein signature"
FT	Domain	60..144
FT	Modified-site	/note= "similarity to EF-hand calcium binding domain"
FT	Modified-site	69
FT	Modified-site	/note= "potential N-glycosylation"
FT	Modified-site	71
FT	Modified-site	/note= "potential O-phosphorylation"
FT	Modified-site	108
FT	Modified-site	/note= "potential O-phosphorylation"
FT	Region	115..128
FT	Region	/note= "similarity to annexin type I signature"
XX		
XX	MO200005368-A1.	
XX	03-FEB-2000.	
XX	19-JUL-1999;	99WO-US012385.
XX	20-JUL-1998;	98US-0160076P.

XX (INCY-) INCYTE PHARM INC.
 XX Bandman O, Tang YT, Corley NC, Guegler KU, Baughn MR, Junning Y;
 XX WPI; 2000-182695/16.
 XX N-PSDB; AA58233.
 XX New regulatory proteins and polynucleotides useful for treating and
 PT detecting neurological disorders including epilepsy, Alzheimer's disease,
 PT and stroke and cardiovascular disorders e.g. cardiomyopathy, myocarditis,
 PT tachyarrhythmia.
 PS Claim 1; Fig 1A-B; 74pp; English.
 XX The present sequence is that of novel human calcium regulatory protein 1
 CC (CarBG-1), as deduced from a consensus sequence (see AA58233) of
 CC isolated cDNA clones. The protein has chemical and structural similarity
 CC with human S100- α 1. Northern analysis showed that CarBG-1 is expressed
 CC exclusively in lung tissue, including foetal and cancerous lung tissue.
 CC The invention provides CarBG-1 and CarBG-2 polynucleotides and
 CC polypeptides, expression vectors, host cells, antibodies, agonists and
 CC antagonists. The polypeptides, polynucleotides and agonists are useful
 CC for treating or preventing neurological disorders such as epilepsy,
 CC ischaemic cerebrovascular diseases, stroke, Alzheimer's disease etc., and
 CC cardiovascular disorders such as cardiomyopathy, myocarditis, Duchenne's
 CC muscular dystrophy, tachyarrhythmia etc. Antagonists and vectors
 CC comprising antisense sequences are useful for treating and preventing
 CC cancers such as adenocarcinoma, leukaemia, lymphoma and melanoma
 XX Sequence 147 AA;
 SQ
 Query Match 100.0%; Score 147; DB 3; Length 147;
 Best Local Similarity 100.0%; Pred. No. 7.8e-140; Indels 0; Gaps 0;
 Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGCGMSTODKSLHLEGDPPNSAAPSTGAPKPRKISISQKLSVYKLRKCSDEKAI 60
 Db 1 MGCGMSTODKSLHLEGDPPNSAAPSTGAPKPRKISISQKLSVYKLRKCSDEKAI 60
 QY 61 ATTALIFRNSSDSGKLEKAIADLLQTOFRNFAEGQETPKYRELISLDEHTENKIDF 120
 Db 61 ATTALIFRNSSDSGKLEKAIADLLQTOFRNFAEGQETPKYRELISLDEHTENKIDF 120
 QY 121 EDPMILLSTVMSDLLQNIIRNYKIMK 147
 Db 121 EDPMILLSTVMSDLLQNIIRNYKIMK 147
 RESULT 2
 ID AAM40026 standard; protein; 143 AA.
 AC AAM40026;
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 3171.
 XX Human, nocitropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 PN WO20015312-A1.
 PD 26-JUL-2001.
 PF 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.
 XX 21-JAN-2000; 2000US-00488725.
 XX 25-APR-2000; 2000US-00552317.
 XX 20-JUN-2000; 2000US-00598042.
 XX 19-JUL-2000; 2000US-00620312.
 XX 03-AUG-2000; 2000US-00653450.
 XX 14-SEP-2000; 2000US-00662191.
 XX 19-OCT-2000; 2000US-00693036.
 XX 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 XX N-PSDB; AA159162.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 PS Example 4; SEQ ID NO 3171; 10078pp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA138642-AA42213) with nocitropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX Sequence 143 AA;
 SQ
 Query Match 97.3%; Score 143; DB 4; Length 143;
 Best Local Similarity 100.0%; Pred. No. 8.2e-136; Indels 0; Gaps 0;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 MESTODKSLHLEGDPPNSAAPSTGAPKPRKISISQKLSVYKLRKCSDEKAIATTA 64
 Db 1 MESTODKSLHLEGDPPNSAAPSTGAPKPRKISISQKLSVYKLRKCSDEKAIATTA 60
 QY 65 LIFRNSSDSGKLEKAIADLLQTOFRNFAEGQETPKYRELISLDEHTENKIDFDFPM 124
 Db 61 LIFRNSSDSGKLEKAIADLLQTOFRNFAEGQETPKYRELISLDEHTENKIDFDFPM 120
 QY 125 ILLSTVMSDLLQNIIRNYKIMK 147
 Db 121 ILLSTVMSDLLQNIIRNYKIMK 143
 RESULT 3
 ID AAU16494 standard; protein; 140 AA.
 AC AAU16494;
 DT 07-NOV-2001 (first entry)
 DE Human novel secreted protein. Seq ID 1447.
 XX Human, immunosuppressive; antiarthritic; antiinflammatory; cytostatic;
 KW cardiant; vasotropic; cerebroprotective; nocitropic; neuroprotective;
 KW antibacterial; virucide; fungicide; ophtalmological; vulnerary;

Fri Jun 4 16:15:59 2004

us-09-744-197-1.oligo.rag

Page 3

KM secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KM cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KM cerebral ischaemia; angiogenesis; nervous system disorder;
KM Alzheimer's disease; infection; ocular disorder; corneal infection;
KM wound healing; epithelial cell proliferation; skin ageing; food additive;
KM preservative; antiproliferative.
XX Homo sapiens.
XX MO200155322-A2.
PD 02-AUG-2001.
XX 17-JAN-2001; 2001WC-US001341.
PF 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184668P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209447P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 14-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0218296P.
PR 14-JUL-2000; 2000US-0218496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224516P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225478P.
PR 14-AUG-2000; 2000US-0225587P.
PR 14-AUG-2000; 2000US-0225588P.
PR 18-AUG-2000; 2000US-0226779P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228987P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0230439P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251895P.
PR 08-DEC-2000; 2000US-0251896P.

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251969P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-488783/53.
XX N-PSDB; AAS26481.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11, SEQ ID NO 1447; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 93.9%; Score 138; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 8.8e-131;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 DKSILHEDPNSAPSTGCAPRKPKRISISQSLASVYKLRKCSDEKAIATTAIFPN 69
Db 3 DKSILHEDPNSAPSTGCAPRKPKRISISQSLASVYKLRKCSDEKAIATTAIFPN 62
QY 70 SSDSDGKLEKAIANDLQTOFRNFAEGQETPKYKREILSLDEHTENKIDFEDFKILLIS 129
Db 63 SSDSDGKLEKAIANDLQTOFRNFAEGQETPKYKREILSLDEHTENKIDFEDFKILLIS 122
QY 130 ITVMSDLLONTIRNYKIM 147
Db 123 ITVMSDLLONTIRNYKIM 140

RESULT 4
ABUS5563
ID ABUS5563 standard; protein; 140 AA.
XX
XX ABUS5563;
AC
XX
XX 18-MAR-2003 (first entry)
DT
XX
XX Human novel polypeptide #650.
DE
XX
XX Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW

KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
XX Homo sapiens.
XX
XX US2002132753-A1.
XX
XX 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214885P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226688P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239393P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249239P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI: 2003-147444/14.
XX N-PSDB; ABX73822.

XX 14-SEP-2000.
 XX 09-MAR-2000; 2000WO-US006112.
 PF 11-MAR-1999; 99US-00266513.
 PR 18-AUG-1999; 99US-0149485P.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX Wood M, McGrath A, Shenk MA, Glenn W;
 PI WPI; 2000-579369/54.
 DR
 XX New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide.
 XX
 PS Claim 8; Page 686; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MAOs, homeodomain
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB
 XX
 SQ Sequence 161 AA;
 Query Match 5.4%; Score 8; DB 3; Length 161;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 76 KLEKAIK 83
 DB 146 KLEKAIK 153
 RESULT 7
 AAR9637
 ID AAR9637 standard; protein; 229 AA.
 XX
 AC AAR9637;
 XX
 DT 01-OCT-1996 (first entry)
 XX
 DE Eucalyptus AGE-2 protein.
 XX
 KW Eucalyptus; reproductive gene; AGE2; Arabidopsis; agamous gene;
 KW Antirrhinum; plena gene; sterility; fertility; forest tree;
 KW transgenic plant; flower development; antisense; ribozyme.
 XX
 OS Eucalyptus globulus.
 XX
 PN AU9539013-A.
 XX
 PD 30-MAY-1996.
 XX
 PF 22-NOV-1995; 95AU-00039013.
 XX
 PR 22-NOV-1994; 94AU-00009589.
 XX
 PA (CSTR) COMMONWEALTH SCI & IND RES ORG.
 XX
 PI Harcourt RL, Llewellyn D, Kyoizuka J, Peacock WJ, Southerton S;
 PI Dennis ES;
 XX

DR WPI; 1996-278411/29.
 DR N-PSDB; AAT34432.
 XX
 XX Eucalyptus reproductive genes - useful for prodn. of sterile Eucalyptus
 PT trees useful for establishing wood lot plantations or in re-forestation
 PT projects.
 XX
 PS Disclosure; Page 38-39; 60pp; English.
 XX
 CC AGE-2 protein (AAR9637) is the product of a cDNA clone, AGE2 (AAT34432),
 CC isolated from Eucalyptus globulus cDNA on the basis of its homology to
 CC the Arabidopsis agamous (AG) and Antirrhinum plena (PLE) flower
 CC development genes. AGE-2 shows 62% homology to AG and 64% homology to PLE
 CC protein. AGE2 is very highly expressed in anthers and flower buds. A
 CC related protein, AGE-1 (AAR9633), was also identified. Antisense or
 CC ribozyme constructs of AGE, or of PLE2 and SOE genes (see also AAT34426
 CC and AAT34429-31), can be used to produce sterile transgenic Eucalyptus
 CC trees by minimizing inflorescence
 XX
 SQ Sequence 229 AA;
 Query Match 5.4%; Score 8; DB 2; Length 229;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 76 KLEKAIK 83
 DB 146 KLEKAIK 153
 RESULT 8
 AAE25918
 ID AAE25918 standard; protein; 314 AA.
 XX
 AC AAE25918;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Pokeweed pro-PAB-S protein.
 XX
 KW Necrotic effect; transgenic plant; antiviral protein; pokeweed; PAB-S.
 XX
 OS Phytolacca americana.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /note= "N-terminal processed segment"
 FT Peptide 287..314
 FT /note= "C-terminal processed segment"
 XX
 PN WO200233107-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-GB004593.
 XX
 PR 14-OCT-2000; 2000GB-00025217.
 XX
 PA (ADTE-) ADVANCED TECHNOLOGIES CAMBRIDGE LTD.
 XX
 PI Thomas CRF, McPherson MJ, Atkinson HU, Neelam A;
 PI WPI; 2002-489891/52.
 DR N-PSDB; AAD42715.
 XX
 XX Inducing necrotic effect in specific cells of plant by transforming plant
 PT with a chimeric gene encoding pokeweed antiviral protein and a promoter
 PT which acts in response to application of specific stimulus to plant.
 XX
 PS Claim 24; Page 74-75; 87pp; English.
 XX
 CC The invention relates to a method of inducing a necrotic effect in
 CC specific cells of a plant. The method involves transforming the plant

PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147415P.
 PR 09-AUG-1999; 99US-0147495P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 23-AUG-1999; 99US-0150566P.
 PR 25-AUG-1999; 99US-0150884P.
 PR 26-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151388P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152633P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154439P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157753P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 18-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160982P.
 PR 22-OCT-1999; 99US-0160983P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161921P.
 PR 28-OCT-1999; 99US-0161922P.

PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.
 Query Match 5.4%; Score 8; DB 3; Length 572;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 70 SSDDGKL 77
 Db 383 SSDDGKL 390
 RESULT 10
 ID ABB59387 standard; protein; 1577 AA.
 XX
 AC ABB59387;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 4953.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmacological.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PP 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li FMD, Myers EW,
 XX
 DR WPI; 2001-6556860/75.
 XX
 DR N-PSDB; ABL03490.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 4953; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 1577 AA:
 Query Match 5.4%; Score 8; DB 4; Length 1577;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 38 ISISKOLA 45
 Db 1409 ISISKOLA 1416
 RESULT 11
 ABB42868

ID ABB42868 standard; peptide; 41 AA.
 XX
 AC ABB42868;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #10374 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.
 XX
 PS Claim 27; SEQ ID NO 35503; 639pp + Sequence Listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 41 AA;
 XX
 Query Match 4.8%; Score 7; DB 4; Length 41;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22 SAAPTST 28
 DB 24 SAAPTST 30
 XX
 RESULT 12
 ID AAM36682
 ID AAM36682 standard; protein; 41 AA.
 XX
 AC AAM36682;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #10719 encoded by probe for measuring placental gene expression.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.

XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000663.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX
 PS Claim 27; SEQ ID NO 36951; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see AA131315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 CC
 SQ Sequence 41 AA;
 XX
 Query Match 4.8%; Score 7; DB 4; Length 41;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22 SAAPTST 28
 DB 24 SAAPTST 30
 XX
 RESULT 13
 ID ABB26140
 ID ABB26140 standard; protein; 41 AA.
 XX
 AC ABB26140;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Protein #8139 encoded by probe for measuring heart cell gene expression.
 XX
 KW Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000666.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts.
 XX Claim 15; SEQ ID NO 27910; 530bp; English.
 XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA11535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 41 AA;
 Query Match 4.8%; Score 7; DB 4; Length 41;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22 SAAPTST 28
 Db 24 SAAPTST 30
 RESULT 14
 AAM76575
 ID AAM76575 standard; protein; 41 AA.
 XX AAM76575;
 AC
 XX 06-NOV-2001 (first entry)
 DT
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 36881.
 DE
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukaemia; lymphoma; myeloma.
 OS Homo sapiens.
 XX
 XX W0200157276-A2.
 EN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000668.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 XX Example 4; SEQ ID NO 36881; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention
 XX
 SQ Sequence 41 AA;
 Query Match 4.8%; Score 7; DB 4; Length 41;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22 SAAPTST 28
 Db 24 SAAPTST 30
 RESULT 15
 AAM63761
 ID AAM63761 standard; protein; 41 AA.
 XX AAM63761;
 AC
 XX 05-NOV-2001 (first entry)
 DT
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 35865.
 DE
 XX Human; brain expressed exon; gene expression analysis; probe; microarray;
 KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
 OS Homo sapiens.
 XX
 XX W0200157275-A2.
 EN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000667.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483446/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX Example 4; SEQ ID NO 35866; 650bp + Sequence Listing; English.
 PS
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention
 XX
 SQ Sequence 41 AA;
 Query Match 4.8%; Score 7; DB 4; Length 41;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Fri Jun 4 16:15:59 2004

us-09-744-197-1.oligo.rag

Page 11

Cy	22	SAAPTST	28
Db	24	SAAPTST	30

Search completed: June 3, 2004, 13:28:59
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:26:57 ; Search time 21 Seconds
(without alignments)
673.341 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 147

Sequence: 1 MGCGHSTQDKSLHLEGDPN.....LSITWMDLLQNIINRYKIMK 147

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

1: PIR78:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5.4	433	2	S20963	homeotic protein H
2	5.4	572	2	B84648	hypothetical prote
3	5.4	1051	2	T18351	lmp1 protein - Myc
4	5.4	1365	2	T10822	lmp1 protein - Myc
5	4.8	111	2	AC2497	hypothetical prote
6	4.8	161	2	S3680	GTP-binding membra
7	4.8	183	2	A53917	ADP-ribosylation f
8	4.8	185	2	D70437	ATP synthase F0 su
9	4.8	213	2	AH2636	conserved hypotet
10	4.8	237	2	G97418	hypothetical prote
11	4.8	247	2	S60307	fibp6 protein - gar
12	4.8	259	2	T01700	hypothetical prote
13	4.8	259	2	H84148	transposase (12) B
14	4.8	255	2	H87703	citrate lyase beta
15	4.8	370	2	H63291	conserved hypotet
16	4.8	378	2	C83853	cystathionine gamm
17	4.8	402	2	E86185	hypothetical prote
18	4.8	415	2	T09213	mRNA-binding prote
19	4.8	429	2	AG3585	aldehyde dehydroge
20	4.8	429	2	S46801	hypothetical prote
21	4.8	431	1	W1026	homeotic protein H
22	4.8	451	2	AF0441	probable xanthine/
23	4.8	481	2	F71048	transposase (12) B
24	4.8	504	2	B83772	serotonin receptor
25	4.8	504	2	S27269	serine-rich protei
26	4.8	534	2	T33903	transposase (12) B
27	4.8	552	2	C83965	transposase (12) B
28	4.8	561	2	A84113	transposase (12) B
29	4.8	573	2	S75748	flavoprotein, 63.5

30	7	4.8	588	2	S52524	probable membrane
31	7	4.8	618	2	AD1161	two-component sens
32	7	4.8	618	2	AD1520	conserved hypotet
33	7	4.8	685	2	B69401	gene PC326 protein
34	7	4.8	747	2	S37694	heavy-metal transp
35	7	4.8	790	1	G69071	probable membrane
36	7	4.8	804	2	AG0565	hypothetical prote
37	7	4.8	805	2	A12157	hypothetical prote
38	7	4.8	805	2	T25795	hypothetical prote
39	7	4.8	890	2	H69877	calcium-transport
40	7	4.8	980	2	T24556	hypothetical prote
41	7	4.8	1036	2	AG1326	alpha-mannosidase
42	7	4.8	1807	2	S03124	vitellogenin A2 pr
43	7	4.8	1807	2	T38495	hypothetical prote
44	7	4.8	3005	1	GNVSTV	genome polypeptide
45	6	4.1	16	2	S29631	beta-glucosidase (

ALIGNMENTS

RESULT 1
S20963
homeotic protein Hox B3 - mouse
N/Alternate names: homeotic protein Hox 2.7
C/Species: Mus musculus (house mouse)
C/Date: 22-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 20-Aug-1999
C/Accession: S20963; D42694
R/Shaw, M.H.; Hunt, P.; Nonchev, S.; Papalopulu, N.; Graham, A.; Boncinelli, E.; Krum
EMBO J. 11, 1825-1836, 1992
A/Title: Analysis of the murine Hox-2.7 gene, conserved alternative transcripts with
A/Reference number: S20963; MUID:92256392; PMID:1562411
A/Accession: S20963
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-433 <SHA>
A/Cross-references: GB:X66177; GB:S35628; GB:S35738; NID:G312229; PID:CAA46951.1; PI
R/Nazareli, A.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 2883-2887, 1992
A/Title: Hox-1.11 and Hox-4.9 homeobox genes
A/Reference number: A42694; MUID:92212934; PMID:1348361
A/Accession: D42694
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 213-238 <NAZ>
A/Note: sequence extracted from NCBI backbone (NCBIN:92310, NCBI:92316)
C/Keyfamily: homeotic protein Hox B3; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:192-248/Domain: homeobox homology <Hox>

Query Match 5.4%; Score 8; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 PSAAPTST 28
Db 92 PSAAPTST 99

RESULT 2
B84648
hypothetical protein At2g5420 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: B84648
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Motilal, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Unayama, L.; Talion,
euse, D.; Nierman, W.C.; White, O.; Eisen, D.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: B84648
A/Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-572 <STO>
 A:Cross-references: GB:AE002093; NID:94432854; PIDN:AAD20702.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g25420
 A:Map position: 2

Query Match 5.4%; Score 8; DB 2; Length 572;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 SSDSDGKL 77
 |||||
 DB 383 SSDSDGKL 390

RESULT 3

Imp1 protein - Mycoplasma hominis
 C:Species: Mycoplasma hominis
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
 C:Accession: T18351
 R:Jensen, L.T.; Ladefoged, S.; Birkelund, S.; Christiansen, G.
 A:Title: Selection of Mycoplasma hominis PG21 deletion mutants by cultivation in the pre
 A:Reference number: Z18884; MUID:95369882; PMID:7543881
 A:Accession: T18351
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1051 <JEN>
 A:Cross-references: EMBL:U21961; NID:9790241; PID:9790242; PIDN:AAA81012.1
 C:Genetics:
 A:Gene: Imp1
 A:Genetic code: SGC3

Query Match 5.4%; Score 8; DB 2; Length 1051;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 AIAATTAI 66
 |||||
 DB 20 AIAATTAI 27

RESULT 4

Imp1 protein - Mycoplasma hominis
 C:Species: Mycoplasma hominis
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 07-Dec-1999
 C:Accession: T30822
 R:Jensen, L.T.; Ladefoged, S.; Birkelund, S.; Christiansen, G.
 A:Title: Selection of Mycoplasma hominis PG21 deletion mutants by cultivation in the pre
 A:Reference number: Z18884; MUID:95369882; PMID:7543881
 A:Accession: T30822
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1365 <JEN>
 A:Cross-references: EMBL:U21962; NID:9790243; PID:9790244; PIDN:AAA81013.1
 C:Genetics:
 A:Gene: Imp1
 A:Genetic code: SGC3

Query Match 5.4%; Score 8; DB 2; Length 1365;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 AIAATTAI 66
 |||||
 DB 20 AIAATTAI 27

RESULT 5

AC2497
 Hypothetical protein all7155 [imported] - Noctoc sp. (strain PCC 7120) plasmid pcc712
 C:Species: Noctoc sp. PCC 7120
 A:Note: Noctoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AC2497
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriki
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takarawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 6: 205-213, 2001
 A:Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AC2497
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-111 <KUR>
 A:Cross-references: GB:BA000020; PIDN:BAF78239.1; PID:97135693; GSPDB:GN00180
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all7155
 A:Genome: plasmid

Query Match 4.8%; Score 7; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ILSELDE 112
 |||||
 DB 103 ILSELDE 109

RESULT 6

S36680
 GMP-binding membrane protein lepA - Pseudomonas fluorescens (fragment)
 C:Species: Pseudomonas fluorescens
 C:Date: 22-Jan-1994 #sequence_revision 12-Apr-1996 #text_change 26-Aug-1999
 C:Accession: S36680; S22413
 R:Black, M.T.; Munn, U.G.R.; Allsop, A.S.
 A:Title: On the catalytic mechanism of prokaryotic leader peptidase 1.
 A:Reference number: S22413; MUID:92189595; PMID:1546969
 A:Accession: S22413
 A:Molecule type: DNA
 A:Residues: 1-161 <BLA1>
 A:Cross-references: EMBL:X56466; NID:945509; PIDN:CAA39638.1; PID:945510
 R:Black, M.T.; Munn, U.G.R.; Allsop, A.S.
 Biochem. J. 282, 539-543, 1992
 A:Title: On the catalytic mechanism of prokaryotic leader peptidase 1.
 A:Reference number: S22413; MUID:92189595; PMID:1546969
 A:Accession: S22413
 A:Molecule type: DNA
 A:Residues: 125-161 <BLA2>
 A:Cross-references: EMBL:X56466
 C:Genetics:
 A:Gene: lepA
 C:Suprafamily: GTP binding; membrane protein

Query Match 4.8%; Score 7; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 SVKALRK 52
 |||||
 DB 106 SVKALRK 112

RESULT 7

A53917
 ADP-ribosylation factor 3 - yeast (Saccharomyces cerevisiae)
 A:Alternate names: protein O3172; protein YOR094w; protein YOR3172w
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-Oct-1994 #sequence_revision 10-Feb-1995 #text_change 02-Feb-2001
 C:Accession: A53917; S61655; S66979
 R:Lee, F.J.S.; Stevens, L.A.; Kao, Y.L.; Moss, J.; Vaughan, M.

J. Biol. Chem. 269, 20931-20937, 1994
 A>Title: Characterization of a glucose-repressible ADP-ribosylation factor 3 (ARF3) from
 A:Reference number: A53917; PMID:9434251; PMID:8063710
 A:Accession: A53917
 A:Molecule type: mRNA
 A:Residues: 1-183 <EB>
 A:Cross-references: GB:128996; NID:g576548; PIDN:AA61614.1; PID:g576549
 R:Benes, V.; Andrade, M.A.; Reckmann, S.; Teodoru, C.; Banreivi, A.; Sander, C.; Valencia
 R:Benes, V.; Andrade, M.A.; Reckmann, S.; Teodoru, C.; Banreivi, A.; Sander, C.; Valencia
 Submitted to the EMBL Data Library, December 1995
 A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
 A:Reference number: S61643
 A:Accession: S61655
 A:Molecule type: DNA
 A:Residues: 1-183 <BN>
 A:Cross-references: EMBL:X94335; NID:g1262139; PIDN:CAA64016.1; PID:g1164941
 R:Yoss, H.; Benes, V.; Reckmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansoorge, W.
 Submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66965
 A:Accession: S66979
 A:Molecule type: DNA
 A:Residues: 1-183 <YOS>
 A:Cross-references: EMBL:Z75002; NID:g1420266; PIDN:CAA92921.1; PID:g1420267; MIPS:YOR09
 A:Experimental source: Strain S288C
 A:Gene: SGD:ARF3
 C:Genetics:
 A:Map position: 15R
 A:Superfamily: ADP-ribosylation factor
 C:Keywords: blocked amino end, lipoprotein, myristylation, nucleotide binding, P-loop
 F:24-31/Region: nucleotide-binding motif A (P-loop)
 F:89-94/Region: nucleotide-binding motif B
 F:126-129/Region: GTP-binding NKXD motif
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F:3/Modified site: aspartic acid (Asn) #status predicted

Query Match 4.8%; Score 7; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 ATTALIF 67
 |||||
 DB 84 ATTALIF 90

RESULT 8
 D70437
 ATP synthase F0 subunit b - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C:Accession: D70437
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy
 V.
 Nature 392, 353-358, 1998
 A>Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; PMID:9819666; PMID:9537320
 A:Accession: D70437
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-185 <MOF>
 A:Cross-references: GB:AE000746; NID:g2983925; PIDN:AA07478.1; PID:g2983929; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: acpF2

Query Match 4.8%; Score 7; DB 2; Length 185;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 IAKDLLO 87
 |||||
 DB 153 IAKDLLO 159

RESULT 9
 AH2636
 conserved hypothetical protein Atu0491 [imported] - Agrobacterium tumefaciens (strain
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AH2636
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
 i, Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
 ster, E.W.
 A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:21608550; PMID:11743193
 A:Accession: AH2636
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-213 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AA41510.1; PID:g17738838; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0491
 A:Map position: circular chromosome

Query Match 4.8%; Score 7; DB 2; Length 213;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 AIAATTL 65
 |||||
 DB 98 AIAATTL 104

RESULT 10
 G97418
 hypothetical protein AGR_C 869 [imported] - Agrobacterium tumefaciens (strain C58, Ce
 C:Species: Agrobacterium tumefaciens
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: G97418
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wolfram, C.; Allinger, M.; Doughy, D.; Scott, C.; Laprae, C.; Markelz,
 Science 294, 2323-2328, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:21608551; PMID:11743194
 A:Accession: G97418
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-237 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86304.1; PID:g15155416; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C 869
 A:Map position: circular chromosome

Query Match 4.8%; Score 7; DB 2; Length 237;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 AIAATTL 65
 |||||
 DB 122 AIAATTL 128

RESULT 11
 S60307
 fbp6 protein - garden petunia
 C:Species: Petunia x hybrida (garden petunia)
 C>Date: 12-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 24-Sep-1999
 C:Accession: S60307
 R:Angenent, G.C.; Franken, J.; Busecher, M.; Colombo, L.; van Tunen, A.J.
 Plant J. 4, 101-112, 1993
 A>Title: Petal and stamen formation in petunia is regulated by the homeotic gene fbp1
 A:Reference number: S60307; PMID:94035167; PMID:8106081
 A:Accession: S60307

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-247 <ANG>
 A:Cross-references: EMBL:X68675; NID:g396198; PIDN:CAA48635.1; PID:g396199
 C:Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
 F:18-73/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GKEKAI 81
 DB 145 GKEKAI 151

RESULT 12
 T01700
 hypothetical protein - maize
 C:Species: Zea mays (maize)
 C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 17-Nov-2000
 C:Accession: T01700; T04388
 R:Menar, M.; Ambrose, B.; Mealey, R.B.; Briggs, S.P.; Yanofsky, M.F.; Schmidt, R.J.
 unpublished results 1997, cited by EMBL
 A:Reference number: Z14402
 A:Accession: T01700
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-259 <MEN>
 A:Cross-references: EMBL:L81162; NID:g2529339; PIDN:AAB81103.1; PID:g2529340
 A:Experimental source: strain A632
 R:Thiessen, G.; Strater, T.; Fischer, A.; Seadler, H.
 Gene 156, 155-166, 1995
 A:Title: Structural characterization, chromosomal localization and phylogenetic evaluat
 A:Reference number: Z15328; MUID:95278740; PMID:7758952
 A:Accession: T04388
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 46-47, 'S', 49-259 <THE>
 A:Cross-references: EMBL:X81200; NID:g854645; PIDN:CAA57074.1; PID:g106768
 C:Genetics:
 A:Gene: ZMM2
 F:34-89/Domain: serum response factor squa; serum response factor DNA-binding domain homol

Query Match
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GKEKAI 81
 DB 162 GKEKAI 166

RESULT 13
 H84148
 transposase (12) BH3992 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: H84148
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: H84148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-259 <STO>
 A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BA07711.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3992

Query Match
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LSELDH 113
 DB 10 LSELDH 16

RESULT 14
 A87703
 citrate lyase beta subunit, probable [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 24-Nov-2003
 C:Accession: A87703
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kc
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: A87703
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-295 <STO>
 A:Cross-references: GB:AE005673; NID:g13425417; PIDN:AAK25621.1; GSPDB:GN0148
 C:Genetics:
 A:Gene: CC3655
 C:Superfamily: citrate lyase, subunit beta

Query Match
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 SNAFST 28
 DB 115 SNAFST 121

RESULT 15
 H69291
 conserved hypothetical protein AF0336 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: H69291
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Doc
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirtness, E
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arc
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: H69291
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-370 <KLE>
 A:Cross-references: GB:AE001081; GB:AE000782; NID:g2689404; PIDN:AAB90898.1; PID:g26;

Query Match
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 REILSEL 110
 DB 270 REILSEL 276

Search completed: June 3, 2004, 13:30:55
 Job time : 22 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	541	73.0	147	11	Q8C9X1	Q8C9X1 mus musculi
2	133.5	18.0	94	11	Q9D3M4	Q9D3M4 mus musculi
3	133.5	18.0	94	11	Q91V77	Q91V77 m 11 days
4	122.5	16.5	101	13	Q93395	Q93395 balvelinus
5	121.5	16.4	79	11	Q9J088	Q9J088 mus musculi
6	112.5	15.2	99	13	Q8AYJ2	Q8AYJ2 squallus ac
7	108.5	14.6	100	13	Q7ZVU4	Q7ZVU4 brachydanio
8	103.5	14.0	92	11	Q9Z5T3	Q9Z5T3 cricetus
9	103.5	14.0	101	6	Q9TV5T3	Q9TV5T3 canis fami1
10	89.5	12.1	546	10	Q7XU87	Q7XU87 cryza sativ
11	86.5	11.7	574	11	Q8C9P5	Q8C9P5 mus musculi
12	86.5	11.7	730	11	Q8BRV2	Q8BRV2 mus musculi
13	86.5	11.7	847	11	Q7R585	Q7R585 mus musculi
14	86	11.6	441	2	Q7XZQ3	Q7XZQ3 mycoplasma
15	85.5	11.5	213	4	Q9H4U1	Q9H4U1 homo sapien
16	85.5	11.5	1037	3	Q9P3K1	Q9P3K1 neurospora

Q7FWR2 mycoplasma
Q862H7 bos taurus
Q8AM5 plasmidum
Q8ZK3 caenorhabdi
Q21081 caenorhabdi
Q80Y4 mus musculu
Q7TO6 mus musculu
Q601177 schizosacch
Q8UX6 gallus gall
Q17602 caenorhabdi
Q9Y6Y7 homo sapien
Q8LEW0 arabidopsis
Q8JLZ4 arabidopsis
Q8FW3
Q949U0 arabidopsis
Q2ZSA2 arabidopsis
Q8YXU4
Q2Z4X3 streptomyce
Q8NSG2
Q8NYG6 homo sapien
Q8NA35 caenorhabdi
Q8ISFP5 caenorhabdi
Q8ISFP6 caenorhabdi
Q8ISF7 caenorhabdi
Q8SE25 oryza sativ
Q8BLX1 mus musculu
Q8VEE7 arabidopsis
Q81390 nicotiana t

ALIGNMENTS

RESULT 1	
Q8C9X1	
ID Q8C9X1	PRELIMINARY;
	PRT; 147 AA

DT_01-MAR-2003 (TREMblrel_23, Created)
 DT_01-MAR-2003 (TREMblrel_23, Last sequence update)
 DT_01-OCT-2003 (TREMblrel_25, Last annotation update)
 DE_Hypothetical_EF-hand_containing_protein.
 DE_Mus_musculus (Mouse).
 OC_Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC_Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
 NCBI_TaxId=10090;
 NCBI_TaxId=10090;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

KL Nucleic acid binding site 420;563-757100264
 DR EMBL; AK040281; BAC30559.1; -
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR InterPro; IPR002048; EF-hand.
 KW Hypothetical protein.
 SQ SEQUENCE 147 AA; 16415 MW; F1CE14DBFB82ED59 CRC64

Query Match	73.0%;	Score 541;	DB 11;	Length 147;
Best Local Similarity	72.1%;	Pred. No. 4.9e-42;		
Matches 106;	Conservative 16;	Mismatches 25;	Indels 0;	Gaps 0

[illegible]

```

QY 121 EDPMTLISITVMSDDLQNTKRYKMK 147
DB 121 EDPVLLISLAIMSDDLNNMENTMK 147

RESULT 2
Q9D3M4 PRELIMINARY; PRT; 94 AA.
AC Q9D3M4 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE S100 calcium binding protein A1.
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Suzuki R., Tomita M., Wagner L.,
RA Schmitt L.M., Straub F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barab G.,
RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein N.C., But C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinchik S., Hill D., Hotmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli U., Mombaefts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wymshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1 SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL: AK017279; BAB30670.1; -.
DR HSSP: P04631; 1B4C.
DR MGD: MGI:1338917; S100a1.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR InterPro: IPR001751; CAPP_S100.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; EFhand; 1.
DR Pfam: PF01023; S_100; 1.
DR ProDom: PD003407; CAPP_S100; 1.
DR ProDom: PD000012; EF-hand; 1.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100 CAPP; 1.
SQ SEQUENCE 94 AA; 10487 MW; 7C8E817D135C2D15 CRC64;

Query Match 18.0%; Score 133.5; DB 11; Length 94;
Best Local Similarity 34.1%; Pred. No. 8.2e-05;
Matches 28; Conservative 22; Mismatches 29; Indels 3; Gaps 1;

QY 54 SDLEKATATATLIPNSDSDG--KLEKAKDLQCPNFAEGQETPKREIISL 110
DB 3 SFLSMEMLIVFVFAHSGKEDKRYKSKELNQLQTELAGFDVQKADAVDKVTEL 62
QY 111 DHTENKADDFEDFILLISITY 132
DB 63 DENGGEVDFEKEYVVALVALTY 84

RESULT 3
Q91V77 PRELIMINARY; PRT; 94 AA.
AC Q91V77;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE 11 days embryo cDNA, RIKEN full-length enriched library,
DE clone:270008BD09, full insert sequence (S100 calcium binding protein
DE A1) (Adult male kidney cDNA, RIKEN full-length enriched library,
DE clone:0610031F03, full insert sequence).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Arakawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,
RA Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,
RA Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,
RA Shibata K., Shibata Y., Shingawa A., Shiraki T., Sogabe Y.,
RA Suzuki H., Tagami T., Tagawa A., Takahashi F., Tanaka T., Tejima Y.,
RA Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Kono H., Akiyama U., Nishi K., Kikunai T., Tachio H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto K., Macnuto H., Sakaguchi S., Ikegami T., Kashinaga K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuyama S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RA Strusberg R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
[7]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Du X.-J., Cole T.J., Tennis N., Gao X.-M., Kontgen F., Kemp B.E.,

```

```

Matches 24; Conservative 28; Mismatches 40; Indels 3; Gaps
Qy 54 SDLEKAIATTAALIFENSSDSDG--KLEKAIAKDLQTOFNFAGGQETPKREILSEL 110
Db 3 SDLESSEMLIVFHRVADKDDCNLTSLKEKELKEMOTELASFLKSKDPDAIDTIVKDL 62
Qy 111 DEHTENKLDPFEDFMILLISTVMSDDLGNINRWKI 145
Db 63 DQNGDGKVSFEFEVSLVWGLSIACBOIYQIATKCV 97

RESULT 5
QJUL08
ID QJUL08 PRELIMINARY; PRT; 79 AA.
AC QJUL08;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE S100 calcium binding protein A1 (Fragment).
GN S100A1.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiewitz R., Lyons G.R., Schafer B.W., Heizmann C.W.;
RT "transcriptional regulation of S100A1 and expression during mouse
RT heart development.";
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY; BELONGS TO THE S-100 FAMILY.
DR EMBL; AF218153; AAF32320.1; -.
DR HSSP; P04631; 1B4C.
DR MGD; MGI:1338917; S100a1.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; Cabp S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S_100; 1.
DR Pfam; PF00036; ehanda; 1.
DR ProDom; PD003407; Cabp_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON TER
FT 79
FT 79
SQ SEQUENCE 79 AA; 8863 MW; F94EDA3A798615D5 CRC64;

Query Match 16.4%; Score 121.5; DB 11; Length 79;
Best Local Similarity 33.8%; Pred. No. 0.00084;
Matches 26; Conservative 20; Mismatches 28; Indels 3; Gaps 1
Qy 54 SDLEKAIATTAALIFENSSDSDG--KLEKAIAKDLQTOFNFAGGQETPKREILSEL 110
Db 3 SDLESSEMLIVFHRVADKDDCNLTSLKEKELKEMOTELASFLKSKDPDAIDTIVKDL 62
Qy 111 DEHTENKLDPFEDFMILLISTVMSDDLGNINRWKI 127
Db 63 DQNGDGKVSFEFEVSLVWGLSIACBOIYQIATKCV 79

RESULT 6
Q8AVJ2
ID Q8AVJ2 PRELIMINARY; PRT; 99 AA.
AC Q8AVJ2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE S-100 calcium-binding protein A1.
OS Squalus acanthias (Spiny dogfish).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
CC Elasmobranchii; Squala; Squalidae; Squalidae; Squalus.
RN [1]
RP SEQUENCE FROM N.A.

```


GN MTS1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madin-Darby; TISSUE=Kidney;
 RA Miyamoto H., Hasegawa K., Kim K., Sato H.;
 RT "Expression of metastasis associated metal gene is co-induced with
 membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
 transformation and tubular formation of madin darby canine kidney
 RT (MDCK) epithelial cells."
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AB031064; BAA83419.1; -.
 DR HSSP; P30801; 1A03.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; EFhand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR ProDom; PD003407; CAP_S100; 1.
 DR ProDom; PD000012; EF-hand; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CAP; 1.
 SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;

Query Match 14.0%; Score 103.5; DB 6; Length 101;
 Best Local Similarity 25.9%; Pred. No. 0.051;
 Matches 21; Conservative 24; Mismatches 33; Indels 3; Gaps 1;

QY 56 LKATATTTLLFRNSDSDG---KLEKAIKDLQTFQFNFAEGQETPKYREILSEIDE 112
 DB 5 LKALDVMVSTFKYSGKSGKGFNKLRLKRLRLPSPLGKRTDRAQKMSNDS 64
 QY 113 HTENKIDFEDFMILLSTITVM 133
 DB 65 NRDNVDFOEYCVFLSCVAMM 85

RESULT 10

Q7XU87 PRELIMINARY; PRT; 846 AA.

AC Q7XU87
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE OSUNBA0029H02.25.
 GN OSUNBA0029H02.25.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriophytidae; Oryzaceae; Oryza.
 NCBI_TaxID=4530;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
 RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
 RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
 RA Chen L., Fan D.L., Deng Q.J., Zhang L., Lu Y.Q., Xu S.L., Liu X.H.,
 RA Lu T.T., Zhang Y.J., Wang Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
 RA Qian Y.M., Yang K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
 RA Zhang R.Q., Guan J.P., Hong G.F.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL065594; CAD41490.1; -.

SQ SEQUENCE 846 AA; 92507 MW; 1AE0398FF57CE402 CRC64;
 Query Match 12.1%; Score 89.5; DB 10; Length 846;
 Best Local Similarity 25.2%; Pred. No. 12;
 Matches 33; Conservative 22; Mismatches 41; Indels 35; Gaps 6;

QY 7 STQDKSLHLEDP-----NPSAPSTCAPKPKRISIKOLASVKALRKSQDLKKA 59
 DB 436 ADDSKSHVDTNPAGKRGKRRPPS-----KSHKKNGK-----GKVGSLSK 479
 QY 60 IATTAIFRNSDSDGKLEKAIK-DLIQTFQFNFAEGQETPKYREILSEL-----DEH 113
 DB 480 KADAV-----SDSGRATRIAKDDIKSFYKGTGEGSSKXKXENLKQEDTPPDD 533
 QY 114 TENKIDFEDFM 124
 DB 534 TDEDLSLKDTIV 544

RESULT 11

08CCF5 PRELIMINARY; PRT; 574 AA.

ID 08CCF5
 AC 08CCF5
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Vav 3 oncogene.
 GN VAV3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA the FANTOM Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK033253; BAC28212.1; -.

DR MGD; MGI:1868518; Vav3.
 DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldenhyde-dehydr.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR003247; CH type.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR001331; DAG_CD24.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000219; RHOGEF.
 DR InterPro; IPR003096; SMW2_calponin.
 DR Pfam; PF00307; CH; 1.
 DR Pfam; PF00130; DAG_PE-bind; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00621; RHOGEF; 1.
 DR PRINTS; PR00688; SMW2CALPONIN.
 DR ProDom; PD001527; CH type; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00033; CH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RHOGEF; 1.
 DR PROSITE; PS00707; ALDENHYDE_DEHYDR_CYS; 1.
 DR PROSITE; PS50021; CH; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS00741; PH; 1.
 DR PROSITE; PS50010; PH_DOMAIN; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 SQ SEQUENCE 574 AA; 66836 MW; 7947CEB3543B6F2B CRC64;

Query Match 11.7%; Score 86.5; DB 11; Length 574;
 Best Local Similarity 28.7%; Pred. No. 15;
 Matches 29; Conservative 14; Mismatches 39; Indels 19; Gaps 3;

QY 53 CSDELKAIATATLIFNNSSDSGKLE-----KAIKDLLOTOFNFAGQETKRY 103
 DB 282 CSGVESALISNDYISKTEKEDVKLKEBSCSRANKGFTLRDLVVPMQRYL-----KY 334
 QY 104 REITSELDEHTENKLPEDFMILLISITVMSDLLONIRVK 144
 DB 335 HLLQELVKTHTDPEMEKAN---LKLALDAMKDLAQYVNEVK 372
 RESULT 12
 ID Q8BRV2 PRELIMINARY; PRT; 730 AA.
 AC Q8BRV2;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Vav 3 oncogene (Fragment).
 GN VAV3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
 RX MEDLINE=22354683; PubMed=12466851;
 RA THE PANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK041249; BAC30879.1; -
 DR PIR; P70558; P70677.
 DR MGI; 1868518; Vav3.
 DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldenhyde dehydr.
 DR InterPro; IPR003447; CH type.
 DR InterPro; IPR002219; DAG PE-bind.
 DR InterPro; IPR001331; GDS_CDC24.
 DR InterPro; IPR001849; PA.
 DR InterPro; IPR000219; RHOGEF.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00307; CH; 1.
 DR Pfam; PF00130; DAG PE-bind; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00621; RHOGEF; 1.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00688; SM22CALPONIN.
 DR PRODOM; PD001527; CH type; 1.
 DR PRODOM; PD000933; SH2; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00033; CH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RHOGEF; 1.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE; PS00021; CH; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS00061; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS00741; DH_1; 1.
 DR PROSITE; PS00109; DH_2; 1.
 DR PROSITE; PS00003; PH_DOMAIN; 1.
 DR PROSITE; PS00001; SH2; 1.
 DR PROSITE; PS00002; SH3; 1.
 DR NON_TER 730
 FT SEQUENCE 730 AA; 84432 MW; 63A9E137D32C447B CRC64;

Query Match 11.7%; Score 86.5; DB 11; Length 847;
 Best Local Similarity 28.7%; Pred. No. 23;
 Matches 29; Conservative 14; Mismatches 39; Indels 19; Gaps 3;
 QY 53 CSDELKAIATATLIFNNSSDSGKLE-----KAIKDLLOTOFNFAGQETKRY 103
 DB 282 CSGVESALISNDYISKTEKEDVKLKEBSCSRANKGFTLRDLVVPMQRYL-----KY 334
 QY 104 REITSELDEHTENKLPEDFMILLISITVMSDLLONIRVK 144
 DB 335 HLLQELVKTHTDPEMEKAN---LKLALDAMKDLAQYVNEVK 372
 RESULT 13
 ID Q7NS85 PRELIMINARY; PRT; 847 AA.
 AC Q7NS85;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Vav 3 oncogene.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=22368257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditchenko L., Marasina K., Parter A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonardo M.J., Cavaletto T.L., Schetz T.F.,
 RA Brownstein W.J., Udell T.B., Tohyuki S., Carninci P., Prange C.J.,
 RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Woley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Kravitz M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RA Strauberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC052739; AAH52739.1; -
 SQ SEQUENCE 847 AA; 97967 MW; 727C9BF50DF8CF19 CRC64;
 Query Match 11.7%; Score 86.5; DB 11; Length 847;
 Best Local Similarity 28.7%; Pred. No. 23;
 Matches 29; Conservative 14; Mismatches 39; Indels 19; Gaps 3;
 QY 53 CSDELKAIATATLIFNNSSDSGKLE-----KAIKDLLOTOFNFAGQETKRY 103
 DB 282 CSGVESALISNDYISKTEKEDVKLKEBSCSRANKGFTLRDLVVPMQRYL-----KY 334
 QY 104 REITSELDEHTENKLPEDFMILLISITVMSDLLONIRVK 144
 DB 335 HLLQELVKTHTDPEMEKAN---LKLALDAMKDLAQYVNEVK 372
 RESULT 14
 ID Q7X2Q3 PRELIMINARY; PRT; 441 AA.

AC 07X203;
 DT 01-OCT-2003 (TRENBLREL. 25, Created)
 DT 01-OCT-2003 (TRENBLREL. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE AVG.
 GN AVG.
 OS Mycoplasma agalactiae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplaemataceae; Mycoplasma.
 CX NCBI_TaxID=2110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Variant 627#3;
 RX MEDLINE=22703458; PubMed=12819065;
 RA Flitman-Tene R., Madañi-Orenstein S., Levisohn S., Yegor D.;
 RT "Variable Lipoprotein Genes of Mycoplasma agalactiae Are Activated in
 RT Vivo by Promoter Addition via Site-Specific DNA Inversions."
 RL Infect. Immun. 71:3821-3830(2003).
 DR EMBL; AY195887; AAC39838.1;
 SQ SEQUENCE 441 AA; 46964 MW; 6D3B960594880B8 CRC64;

Query Match 11.6%; Score 86; DB 2; Length 441;
 Best Local Similarity 24.0%; Pred. No. 12;
 Matches 37; Conservative 27; Mismatches 58; Indels 32; Gaps 6;

QY 19 PNEAAPTSTCAPKMKRISISKOLA-----SVK-----ALRCSDEKAIATTL- 65
 DB 57 PNPGLTPTSTPAKPKGTPEKMAQDTVDSGADISGSIKTPDTTATKSLDKITANNL 116
 QY 66 ---IFRNSSDSDGKLEKAIADLQTFNFAEGQETPKPKREILSELDEHT--ENKLD 120
 DB 117 KVGYSKEDKCKEVEEIKTALV-----AKVTLKGDKLKADLSKNSVVSND 169
 QY 121 EDFMILLISTV-----MSDLLQNTIRNKIMK 147
 DB 170 EGEVTLKEFVETSSGAKAKLSALKITANNL 203

RESULT 15

Q9H4U1 PRELIMINARY; PRT; 213 AA.
 ID Q9H4U1
 AC Q9H4U1;
 DT 01-MAR-2001 (TRENBLREL. 16, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE D14N1.2 (Novel S-100/ICBP type calcium binding domain protein,
 DE similar to trichohyalin) (Fragment).
 GN D14N1.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AL356504; CAC13173.1; -.
 DR HSSP; P02638; ICFP.
 GO; GO:0005509; F:calcium ion binding; IEA.
 DR InterPro; IPR001751; CACP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF01023; S_100; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 FT NON_TER 213
 SQ SEQUENCE 213 AA; 24340 MW; B8C6E0810098E7D2 CRC64;

Query Match 11.5%; Score 85.5; DB 4; Length 213;
 Best Local Similarity 23.2%; Pred. No. 5.6;
 Matches 19; Conservative 22; Mismatches 38; Indels 3; Gaps 1;

DB 2 TDLRSVTVTVIDVYKTKTKDGECCGTSKGEKLEKEHHPVAKNDDDDPTDVIVIMML 61
 QY 111 DEHTENKLDPEDEFMILLISTV 132
 DB 62 DRDHRRLDFTFEFLIMFKLTM 83
 Search completed: June 3, 2004, 13:21:23
 Job time : 39 secs

54 SLEKAIATTLIFRNSSDSGK--LEKAIADLQTFNFAEGQETPKPKREILSEL 110

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 3, 2004, 13:17:56 ; Search time 18 Seconds

(without alignments)
425,240 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 741

Sequence: 1 MGCGMSTQDKSLHLEBDN.....LSITVMDLLGNINVKMK 147

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt 42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138.5	18.7	93	1 S10A_RAT	P35467 ratu
2	133.5	18.0	93	1 S10A_MOUSE	P56565 mus musculu
3	132.5	17.9	93	1 S10A_BOVIN	P02639 bos taurus
4	131.5	17.7	93	1 S10A_HUMAN	P23297 homo sapien
5	123	16.6	97	1 S102_HUMAN	P29034 homo sapien
6	116	15.7	97	1 S102_BOVIN	P10462 bos taurus
7	106.5	14.4	101	1 S104_HUMAN	P26447 homo sapien
8	103.5	14.0	91	1 S10B_MOUSE	P50114 mus musculu
9	102.5	13.8	91	1 S10B_RAT	P04631 ratu
10	102.5	13.8	100	1 S104_MOUSE	P35466 bos taurus
11	102.5	13.8	101	1 S104_BOVIN	P07091 mus musculu
12	102	13.8	122	1 S109_MOUSE	P26783 bos taurus
13	101.5	13.7	91	1 S10B_BOVIN	P02638 bos taurus
14	99.5	13.4	101	1 S104_RAT	P05442 ratu
15	98.5	13.3	91	1 S10B_HUMAN	P04271 homo sapien
16	97.5	13.2	98	1 S102_HUMAN	P04163 sus scrofa
17	96	13.0	95	1 S110_PIG	P08498 homo sapien
18	96	13.0	96	1 S110_HUMAN	P08206 homo sapien
19	96	13.0	101	1 S107_HUMAN	P08050 bos taurus
20	96	13.0	101	1 S107_BOVIN	P08050 bos taurus
21	95.5	12.8	119	1 M126_MOUSE	P08207 mus musculu
22	95	12.8	96	1 S110_MOUSE	P08207 mus musculu
23	94.5	12.8	86	1 S108_RAT	P50115 ratu
24	91.5	12.3	88	1 S108_MOUSE	P27005 mus musculu
25	91	12.3	96	1 S110_CHICK	P27003 gallu
26	89.5	12.1	95	1 S110_HUMAN	P25815 homo sapien
27	89	12.0	94	1 S110_RAT	P06702 homo sapien
28	89	12.0	114	1 S103_HUMAN	P24479 gallu
29	86.5	11.7	101	1 S111_CHICK	P08068 mus musculu
30	86.5	11.7	847	1 VAV3_MOUSE	P02633 bos taurus
31	86	11.6	78	1 S10D_BOVIN	P02632 sus scrofa
32	85	11.5	8545	1 ANCI_CAEEL	
33	84	11.3	78	1 S10D_PIG	

ALIGNMENTS

34	83.5	11.3	843	1 VAV3_RAT	P54100 ratu
35	83.5	11.3	847	1 VAV3_HUMAN	P04044 homo sapien
36	81.5	11.0	98	1 S111_MOUSE	P50543 mus musculu
37	81.5	11.0	779	1 IP2_LISMO	P08746 listeria mo
38	81.5	11.0	782	1 IP2_LISIN	P08746 listeria in
39	81	10.9	100	1 S115_HUMAN	P08746 listeria in
40	81	10.9	112	1 S109_MOUSE	P31725 mus musculu
41	80.5	10.9	1013	1 A60D_DROME	P31927 drosophila
42	80.5	10.9	2493	1 YBA4_YEAST	P35194 saccharomyc
43	80	10.8	1333	1 SMC1_SCHPO	P04383 schizosacch
44	79.5	10.7	93	1 S105_MOUSE	P08845 mus musculu
45	79.5	10.7	99	1 S111_PIG	P31950 sus scrofa

RESULT 1
ID S10A_RAT
AC P35467
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S-100 protein, alpha chain.
GN S100A1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Song W.;
RT Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 10-93 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92076235; PubMed=1742602;
RA Zimmer D.B.; Song W.; Zimmer W.E.;
RT "Isolation of a rat S100 alpha cDNA and distribution of its mRNA in rat tissues."
CC Brain Res. Bull. 27:157-162(1991).
CC - FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC distinct binding sites with different affinities exist for both
CC ions on each monomer. Physiological concentrations of potassium
CC ion antagonize the binding of both divalent cations, especially
CC affecting high-affinity calcium-binding sites.
CC - SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC one alpha and one beta chain.
CC - SUBCELLULAR LOCATION: Cytoplasmic.
CC - TISSUE SPECIFICITY: Although predominant among the water-soluble
CC brain proteins, S-100 is also found in a variety of other tissues.
CC - SIMILARITY: Belongs to the S-100 family.
CC - SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U26358; AAB53657.1; -
CC EMBL: U26357; AAB53657.1; JOINED.
CC EMBL: S68809; AAB20539.2; -
CC PDB: 1K2H; 13-FEB-02.
CC InterPro: IPR001751; CAP S100.
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00036; EF-hand; 1.
CC Pfam: PF01023; S_100; 1.
CC ProDom: PD003407; Cap_S100; 1.

DR PRODOM: PD000012; EF-hand; 1.
 DR PROSITE; PS00018; EF-HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Calcium-binding; zinc; Metal-binding; 3D-structure.
 FT INIT MET 0
 FT CA_BIND 19 32
 FT CA_BIND 2 19
 FT CONFLICT 62 73
 FT CONFLICT 13 13
 FT CONFLICT 56 56
 SO SEQUENCE 93 AA; 10429 MW; D2959A95SEC0651A9 CRC64;
 Query Match 18.7%; Score 138.5; DB 1; Length 93;
 Best Local Similarity 36.6%; Pred. No. 1.8e-05;
 Matches 30; Conservative 21; Mismatches 28; Indels 3; Gaps 1;
 QY 54 SDEKAIATATLAFRNSSDSG--KLEKAIANDLQTOFRNFAEGEETKPKYREILSEL 110
 DB 2 SELESMETLIVNFAHSGEGDKYKSKELKDLQTELSFLDVQADAVDKYKML 61
 QY 111 DEHTENKLPEDFEMILLSTIV 132
 DB 62 DENGDEVDPEFVVAALTV 83
 RESULT 2
 S10A_MOUSE STANDARD; PRT; 93 AA.
 AC P56565; O88949;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE S-100 protein, alpha chain (S100 calcium-binding protein A1).
 GN S100A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
 RA Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
 RA Schellenberg K., Stepien M., Tan F., Underwood K., Moore B.,
 RA Theising B., Wyllie T., Lennon G., Soares B., Wilson R.,
 RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99117144; PubMed=9920416;
 RA Ridinger K., Ilg E.C., Niggli F.K., Heizmann C.W., Schaefer B.W.,
 RT "Clustered organization of S100 genes in human and mouse";
 RL Biochim. Biophys. Acta 1448:254-263(1998).
 CC -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 distinct binding sites with different affinities exist for both
 ions on each monomer. Physiological concentrations of potassium
 ion antagonize the binding of both divalent cations, especially
 affecting high-affinity calcium-binding sites (by similarity).
 CC -!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
 one alpha and one beta chain (by similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AA000715; -; NOT_ANNOTATED_CDS.
 CC EMBL; AA207749; -; NOT_ANNOTATED_CDS.
 CC EMBL; AA500563; -; NOT_ANNOTATED_CDS.
 CC EMBL; AA432539; -; NOT_ANNOTATED_CDS.

DR EMBL; AF087687; AAC64108.1; -.
 DR HSSP; P04631; 1BAC.
 DR MGD; MG11336917; S100a1.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF01023; S100; 1.
 DR PRODOM: PD003407; CABP_S100; 1.
 DR PRODOM: PD000012; EF-hand; 1.
 DR PROSITE; PS000012; EF-hand; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Calcium-binding; zinc; Metal-binding.
 FT INIT MET 0
 FT CA_BIND 19 32
 FT CA_BIND 2 19
 FT CONFLICT 21 21
 FT CONFLICT 21 21
 FT CONFLICT 21 21
 FT CONFLICT 21 21
 SO SEQUENCE 93 AA; 10374 MW; C969E39E1BD705C CRC64;
 Query Match 18.0%; Score 133.5; DB 1; Length 93;
 Best Local Similarity 34.1%; Pred. No. 4.8e-05;
 Matches 28; Conservative 22; Mismatches 29; Indels 3; Gaps 1;
 QY 54 SDEKAIATATLAFRNSSDSG--KLEKAIANDLQTOFRNFAEGEETKPKYREILSEL 110
 DB 2 SELESMETLIVNFAHSGEGDKYKSKELKDLQTELSFLDVQADAVDKYKML 61
 QY 111 DEHTENKLPEDFEMILLSTIV 132
 DB 62 DENGDEVDPEFVVAALTV 83
 RESULT 3
 S10A_BOVIN STANDARD; PRT; 93 AA.
 AC P02639;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE S-100 protein, alpha chain.
 GN S100A1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66248083; PubMed=3755105;
 RA Kuwano R., Meedi T., Usui H., Arai K., Yamakuni T., Ohshima Y.,
 RA Kurihara T., Takahashi Y.,
 RT "Molecular cloning of cDNA of S100 alpha subunit mRNA";
 RL FEBS Lett. 202:97-101(1986).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=81236562; PubMed=7250124;
 RA Isebe T., Okuyama T.,
 RT "The amino-acid sequence of the alpha subunit in bovine brain S-100a
 protein";
 RL Eur. J. Biochem. 116:79-86(1981).
 RN [3]
 RP METAL ION-BINDING PROPERTIES.
 RX MEDLINE=84000339; PubMed=6615778;
 RA Baudier J., Gerard D.,
 RT "Ions binding to S100 proteins: structural changes induced by calcium
 and zinc on S100a and S100b proteins";
 RL Biochemistry 22:3360-3369(1983).
 CC -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 distinct binding sites with different affinities exist for both
 ions on each monomer. Physiological concentrations of potassium

DE S100 calcium-binding protein A2 (S-100L protein) (CAN19).
 GN S100A2 OR S100L.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92196147; PubMed=1372446;
 RA Lee S.W., Tomasetto C., Swisshelm K., Keyomarsi K., Sager R.;
 RT "Down-regulation of a member of the S100 gene family in mammary
 RT carcinoma cells and reexpression by azidothymidine treatment.";
 RU Proc. Natl. Acad. Sci. U.S.A. 89:2504-2508(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96142474; PubMed=9481475;
 RA Wicks R., Franz C., Scholl F.A., Heizmann C.W., Schaefer B.W.;
 RT "Repression of the candidate tumor suppressor gene S100A2 in breast
 RT cancer is mediated by site-specific hypermethylation.";
 RU Cell Calcium 22:243-254(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stephens M., Soares M.B., Bonaldo M.F., Casavant T.E.,
 RA Brownstein W.J., Uscin T.B., Tomshy S., Abramson R.D., Mullany S.J.,
 RA Rada S.S., McQuellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,
 RA Bosak S.A., Loebl P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fabsy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Smalins D.E.,
 RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RU Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 36-39 AND 41-48.
 RX TISSUE=Keratinocytes; PubMed=1286667;
 RA MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.R., Van Damme T., Puype M., Geeser B., Celis J.E.,
 RA Vandekekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RU Electrophoresis 13:960-969(1992).
 CC - FUNCTION: May act as a modulator against excess calcium
 CC accumulation in normal human mammary epithelial cells. May also
 CC play a role in suppressing tumor cell growth.
 CC - TISSUE SPECIFICITY: A subset of epithelial cells amongst which
 CC normal human mammary epithelial cells and keratinocytes.
 CC - DEVELOPMENTAL STAGE: Preferentially expressed in normal human
 CC mammary epithelial cells as opposed to tumor-derived ones. The
 CC level of S100L was shown to correlate inversely with tumor
 CC progression.
 CC - INDUCTION: By growth factors in early G1 phase and probably by
 CC cell-cycle regulation in S phase. DNA methylation probably plays
 CC a direct negative role in suppressing S100L gene expression in
 CC tumor cells.
 CC - MISCELLANEOUS: This protein binds two calcium ions (By
 CC similarity).
 CC - SIMILARITY: Belongs to the S-100 family.
 CC - SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 CC EMBL; M87068; NOT ANNOTATED_CDS.
 CC EMBL; Y07755; CA69033.1; --
 CC EMBL; BC002829; AA02829.2; --
 CC HSSP; P30801; 1A03.
 CC Aarhu/Ghent-2DPAGE; 9027; IEF.
 CC Aarhu/Ghent-2DPAGE; 9051; IEF.
 CC Genew; HGNC:10492; S100A2.
 CC MIM; 176993;
 CC GO; GO:0005509; F:calcium ion binding; NAS.
 CC InterPro; IPR001751; CABP_S100.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF01023; S_100; 1.
 CC ProDom; PD003407; CABP_S100; 1.
 CC ProDom; PD000012; EF-hand; 1.
 CC PROSITE; PS00018; EF_HAND; 1.
 CC PROSITE; PS00303; S100_CABP; 1.
 CC KIM; 176993;
 CC CA_BIND; 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
 CC CA_BIND; 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
 CC CONFLICT; 61 61 S -> N (IN REF. 1).
 CC SEQUENCE; 97 AA; 10986 MW; CFB06CD8DC08D50 CRC64;
 SQ
 Query Match 16.6%; Score 123; DB 1; Length 97;
 Best Local Similarity 31.1%; Pred. No. 0.00041;
 Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;
 QY 53 CSIDLKATATTAIFPNSSDSG--KLEKAIKDLQTFQNFABGQETKRRRLISE 109
 DB 2 CSSLKQALAVLVTFPHKYSQCEDEKPKLKGKELHKLKPLSVGEKVDGKLKMGWS 61
 QY 110 IDEHTENKLPEDFPMILSIYTW-SDLLQ 138
 DB 62 LDENSDQGVDFQRYAVFLALITWCMDFQ 91
 RESULT 6
 ID S102 BOVIN STANDARD; PRT; 97 AA.
 AC P10462;
 DT 01-JUN-1989 (Ref. 11, Created)
 DT 01-JUN-1989 (Ref. 11, Last sequence update)
 DT 28-FEB-2003 (Ref. 41, Last annotation update)
 DE S100 calcium-binding protein A2 (S-100L protein).
 GN Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxId=9913;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX TISSUE=Lung; PubMed=2521861;
 RA MEDLINE=89139574; PubMed=2521861;
 RA Glenney J.R. Jr., Kindy M.S., Zokas L.;
 RT "Isolation of a new member of the S100 protein family: amino acid
 RT sequence, tissue, and subcellular distribution.";
 RU J. Cell Biol. 108:569-578(1989).
 CC - SUBUNIT: Homomultimeric (Probable).
 CC - FUNCTION: Homomultimeric (Probable).
 CC - MISCELLANEOUS: This protein binds two calcium ions.
 CC - SIMILARITY: Belongs to the S-100 family.
 CC - SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC PIR; A30129; A30129.
 CC HSSP; P30801; 1A03.
 CC InterPro; IPR001751; CABP_S100.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF01023; S_100; 1.
 CC ProDom; PD003407; CABP_S100; 1.

DR PRODOM, PD000012, EF-hand; 1.
 DR PROSITE, PS00018, EF_HAND; 1.
 DR PROSITE, PS00303, S100_CABP; 1.
 KW Calcium-binding.
 FT CA BIND 20 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
 FT CA BIND 63 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
 SQ SEQUENCE 97 AA: 10893 MW: 55ACCF60CF9C66 CRC64:
 Query Match 15.7%; Score 116; DB 1; Length 97;
 Best Local Similarity 29.2%; Pred. No. 0.016;
 Matches 26; Conservative 24; Mismatches 35; Indels 4; Gaps 2;
 QY 54 SDLEKAVITLALIFNSSSDSG---KLEKAIKDLLOTFQFNFAEGQETPKYREILSEL 110
 DB 3 SPLEKALVAVMTATFKYSGEGDKFKLSKGEKELHLELPSFQVKEKDESKKMGDL 62
 QY 111 DEHTENKLDPEDEMTLLISTVM-SDLIQ 138
 DB 63 DENSDQVDFOEYAVFALITTMCDPFQ 91
 RESULT 7
 ID S104_HUMAN STANDARD; FRT; 101 AA.
 AC P26447;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Placental calcium-binding protein (Calvesculin) (S100 calcium-binding protein A4) (M1S1 protein).
 GN S100A4 OR CAPL OR M1S1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=93041710; PubMed=1384693;
 RA Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;
 RT "S100 alpha, CAPL, and CACY: molecular cloning and expression
 RL analysis of three calcium-binding proteins from human heart";
 RL Biochemistry 31:10258-10264(1992).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=93342029; PubMed=8341667;
 RA Engelkamp D., Schaefer B., Mattei M.-G., Erne P., Heizmann C.W.;
 RT "Six S100 genes are clustered on human chromosome 1q21:
 RT identification of two genes coding for the two previously unreported
 RT calcium-binding proteins S100D and S100E";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6547-6551(1993).
 RN [3]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=93028421; PubMed=1329089;
 RA Tsuchinsky E.M., Ford H.L., Kramarov D., Reshetnyak E., Grigorian M.,
 RA Zain S., Lukandin E.;
 RT "Transcriptional analysis of the mts1 gene with specific reference to
 RT 5' flanking sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:9146-9150(1992).
 RN [4]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Cervix, and Prostate;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz J.E.,
 RA Brownstein M.J., Udén T.B., Yoshizaki S., Carninci P., Prange C.,
 RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McEwan P.D., McKernan K.J., Malek J.A., Gunaratne P.R.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- SIMILARITY: Belongs to the S-100 family.
 CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, M90563; AA51920.1; -;
 CC EMBL, Z18950; CA879474.1; -;
 CC EMBL, Z33457; CA83880.1; -;
 CC EMBL, BC000838; AA00838.1; -;
 CC EMBL, BC016300; AA016300.1; -;
 CC PIR, A48219; A48219.
 CC PDB, 1M31; 3O-OCT-02.
 CC SWISS-2DPAGE; P26447; HUMAN.
 CC Genew; HGNC:10494; S100A4.
 CC MW: 114210; -;
 CC GO; GO:0005509; P:calcium ion binding; TAS.
 CC InterPro; IPR001751; CAPB_S100.
 CC InterPro; IPR02048; EF-hand.
 CC Pfam; PF000036; ehand; 1.
 CC Pfam; PF01023; S_100; 1.
 CC ProDom; PD003407; CAPB_S100; 1.
 CC ProDom; PD000012; EF-hand; 1.
 CC PROSITE, PS00018; EF_HAND; 1.
 CC PROSITE, PS00303; S100_CABP; 1.
 CC KX Calcium-binding; 3D-structure.
 CC FT CA BIND 20 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
 CC FT CA BIND 63 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
 CC SQ SEQUENCE 101 AA: 11728 MW: 286D2B7B07DB562 CRC64;
 Query Match 14.4%; Score 106.5; DB 1; Length 101;
 Best Local Similarity 28.4%; Pred. No. 0.011;
 Matches 23; Conservative 22; Mismatches 33; Indels 3; Gaps 1;
 QY 56 LKXALVAVMTATFKYSGEGDKFKLSKGEKELHLELPSFQVKEKDESKKMGDL 112
 DB 5 LKXALDVAVSTFKYSGEGDKFKLSKGEKELHLELPSFQVKEKDESKKMGDL 64
 QY 113 HTENKLDPEDEMTLLISTVM 133
 DB 65 NRDNVDVDFQECVFLSCIAMM 85
 RESULT 8
 ID S10B_MOUSE STANDARD; FRT; 91 AA.
 AC P50114;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE S-100 protein, beta chain.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;

RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=BA6/c.
 RX MEDLINE=9338628; PubMed=8376406;
 RA Jiang H., Shah S., Hilt D.C.;
 RT "Organization, sequence, and expression of the murine S100 beta gene.
 RT Transcriptional regulation by cell type-specific cis-acting
 RT regulatory elements.";
 RL J. Biol. Chem. 269:20502-20511(1993).
 CC -1- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 CC distinct binding sites with different affinities exist for both
 CC ions on each monomer. Physiological concentrations of potassium
 CC ion antagonize the binding of both divalent cations, especially
 CC affecting high-affinity calcium-binding sites.
 CC -1- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
 CC one alpha and one beta chain.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Although predominant among the water-soluble
 CC brain proteins, S-100 is also found in a variety of other tissues
 CC (By similarity).
 CC -1- SIMILARITY: Belongs to the S-100 family.
 CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L22144; AAA03075.1; -.
 DR PIR: A48015; A48015.
 DR HESP: P04631; 1B4C.
 DR MGD: MGI:98217; S100b.
 DR GO: GO:0005737; Cytoplasm; NAS.
 DR GO: GO:0005576; Cytosol; NAS.
 DR GO: GO:0005509; F-actin binding; NAS.
 DR GO: GO:0019210; F-kinase inhibitor activity; ISS.
 DR GO: GO:0042803; F-protein homodimerization activity; ISS.
 DR GO: GO:0048155; F-S100 alpha binding; ISS.
 DR GO: GO:0048154; F-S100 beta binding; ISS.
 DR GO: GO:0048156; F-tau protein binding; ISS.
 DR GO: GO:0008270; F-zinc ion binding; ISS.
 DR GO: GO:0048143; F-astrocyte activation; ISS.
 DR GO: GO:0007409; P-axonogenesis; ISS.
 DR GO: GO:0006874; P-calcium ion homeostasis; NAS.
 DR GO: GO:0006112; P-energy reserve metabolism; NAS.
 DR GO: GO:0048151; P-hyperphosphorylation; ISS.
 DR GO: GO:0006917; P-induction of apoptosis; NAS.
 DR GO: GO:0007611; P-learning and/or memory; IMP.
 DR GO: GO:0042035; P-positive regulation of complement activation; ISS.
 DR GO: GO:0048159; P-regulation of cytokine biosynthesis; NAS.
 DR GO: GO:0006417; P-regulation of long-term neuronal synaptic p...; NAS.
 DR InterPro: IPR001751; CAP_S100.
 DR InterPro: IPR02048; EF-hand.
 DR Pfam: PF00036; ehand; 1.
 DR Pfam: PF01023; S-100; 1.
 DR PRODOM: PD003407; CAP_S100; 1.
 DR PROSITE: PS00018; EF_HAND; 1.
 DR PROSITE: PS00303; S100_CAP; 1.
 KW Calcium-binding; Zinc; Metal-binding.
 FT INTR MET 0
 FT CA BIND 18 31 EF-HAND 1 (LOW AFFINITY).
 FT CA BIND 61 72 EF-HAND 2 (HIGH AFFINITY).
 SQ SEQUENCE 91 AA: 10597 NW: 2378AAB8BE1C94D CRC64;
 Query Match 14.0%; Score 103.5; DB 1; Length 91;
 Best Local Similarity 28.4%; Pred. No. 0.018;
 Matches 23; Conservative 21; Mismatches 34; Indels 3; Gaps 1;

Db 1 SELEKAMVALIDVFOYSGREDDKHKSEJKEJLNELSHFLEIRGEVVDKMETL 60
 Cy 111 DEHTENKLPDEDFMILLASIT 131
 Db 61 DEGDCECFQEFMARVAVMT 81
 RESULT 9
 ID S10B RAT STANDARD; PRT; 91 AA.
 AC P04631;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE S-100 protein, beta chain.
 GN S100B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85037924; PubMed=6093041;
 RA Kuwano R., Usui H., Maeda T., Fukui T., Yamanari N., Ohtsuka E.,
 RA Ikahara M., Takahashi Y.;
 RT "Molecular cloning and the complete nucleotide sequence of cDNA to
 RT mRNA for S-100 protein of rat brain."
 RL Nucleic Acids Res. 12:7455-7465(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kuwano R., Usui H., Maeda T., Araki K., Kurihara T., Yamakuni T.,
 RA Ohtsuka E., Ikahara M., Takahashi Y.;
 RT "Molecular cloning and nucleotide sequences of cDNA and genomic DNA
 RT for alpha and beta subunits of S100 protein."
 RL Taniguchi Symp. Brain Sci. 19:243-255(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91359841; PubMed=1653388;
 RA Maeda T., Usui H., Araki K., Kuwano R., Takahashi Y., Suzuki Y.;
 RT "Structure and expression of rat S-100 beta subunit gene."
 RL Brain Res. Mol. Brain Res. 10:199-202(1991).
 RN [4]
 RP SEQUENCE OF 5-91 FROM N.A.
 RX MEDLINE=87137648; PubMed=3818655;
 RA Dunn R., Landry C., O'Hanlon D., Dunn J., Allore R., Brown I.,
 RA Marks A.;
 RT "Reduction in S100 protein beta subunit mRNA in C6 rat glioma cells
 RT following treatment with anti-microtubular drugs."
 RL J. Biol. Chem. 262:3562-3566(1987).
 RN [5]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96387197; PubMed=8794737;
 RA Drohat A.C., Amburgey J.C., Abildgaard F., Starich M.R.,
 RA Baldisseri D.M., Weber D.J.;
 RT "Solution structure of rat apo-S100B(beta beta) as determined by NMR
 RT spectroscopy."
 RL Biochemistry 35:11577-11588(1996).
 RN [6]
 RP STRUCTURE BY NMR.
 RX MEDLINE=98153156; PubMed=9485423;
 RA Drohat A.C., Baldisseri D.M., Rustandi R.R., Weber D.J.;
 RT "Solution structure of calcium-bound rat S100B(beta beta) as
 RT determined by nuclear magnetic resonance spectroscopy."
 RL Biochemistry 37:2729-2740(1998).
 RN [7]
 RP STRUCTURE BY NMR.
 RX MEDLINE=99226808; PubMed=10211826;
 RA Drohat A.C., Tjandra N., Baldisseri D.M., Weber D.J.;
 RT "The use of dipolar couplings for determining the solution structure
 RT of rat apo-S100B."
 RL Protein Sci. 8:800-809(1999).
 CC -1- FUNCTION: Weakly binds calcium but binds zinc very tightly-

distinct binding sites with different affinities exist for both ions on each monomer. Physiological concentrations of potassium ion antagonize the binding of both divalent cations, especially affecting high-affinity calcium-binding sites.

-1- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or one alpha and one beta chain.

-1- SUBCELLULAR LOCATION: Cytoplasmic.

-1- TISSUE SPECIFICITY: Although predominant among the water-soluble brain proteins, S-100 is also found in a variety of other tissues.

-1- SIMILARITY: Belongs to the S-100 family.

-1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X01090; CAA25567.1; -

EMBL; M54919; AAA2096.1; -

EMBL; S53527; -; NOT_ANNOTATED_CDS.

EMBL; S53522; -; NOT_ANNOTATED_CDS.

EMBL; M15705; -; NOT_ANNOTATED_CDS.

PIR; A60046; A26557.

PDB; 1SYM; 07-DEC-96.

PDB; 1OLX; 11-NOV-98.

PDB; 1BAC; 30-DEC-98.

PDB; 1DT7; 26-JUL-00.

PDB; 1WNI; 18-DEC-02.

GO; GO:0005737; Cytoplasm; ISS.

GO; GO:0005766; Cytoplasmic; ISS.

GO; GO:0005809; F-kinase inhibitor activity; ISS.

GO; GO:0019210; F-kinase inhibitor activity; ISS.

GO; GO:0048156; F-kinase activity; ISS.

GO; GO:0008270; F-kinase activity; ISS.

GO; GO:0048143; F-kinase activity; ISS.

GO; GO:0007409; F-kinase activity; ISS.

GO; GO:0006874; F-kinase activity; ISS.

GO; GO:0006112; F-kinase activity; ISS.

GO; GO:0048151; F-kinase activity; ISS.

GO; GO:0006917; F-kinase activity; ISS.

GO; GO:0007611; F-kinase activity; ISS.

GO; GO:0045517; F-kinase activity; ISS.

GO; GO:0042035; F-kinase activity; ISS.

GO; GO:0048169; F-kinase activity; ISS.

GO; GO:0006417; F-kinase activity; ISS.

InterPro; IPR001751; CAP_S100.

InterPro; IPR002048; EF-hand.

Pfam; PF00036; ehand; 1.

Pfam; PF01023; S_100; 1.

ProDom; PD003407; CAP_S100; 1.

ProSite; PS00018; EF_HAND; 1.

ProSite; PS00303; S100 CAP; 1.

KM Calcium-binding; Zinc; Metal-binding; 3D-structure.

CA_BIND 0

INIT MET 0

CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY).

CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).

HELIIX 2 17

TURN 18 19

STRAND 27 27

HELIIX 29 39

TURN 41 42

HELIIX 43 46

HELIIX 50 61

TURN 62 62

STRAND 68 68

HELIIX 70 83

TURN 87 88

SEQUENCE 91 AA; 10613 MW; 2378AAB8BF7134D CRC64;

Query Match 13.8%; Score 102.5; DB 1; Length 91;

Best Local Similarity 28.4%; Pred. No. 0.022; Matches 23; Conservative 21; Mismatches 34; Indels 3; Gaps 1;

Qy 54 SDEKAIATATLTPRSSDS--KLEKAIADLTQTPRNARFQETPKKREILSEL 110
Db 1 SEEEKAWVALIDVFHOYSGREGKHKLSKSELINNELSHLEIKQEVVDKWEIL 60

Qy 111 DEHTENKLDPEDEPMILLISIT 131
Db 61 DEGDGDCDFQEFMAFVSWT 81

RESULT 10

S104 BOVIN STANDARD; PRT; 100 AA.

AC P35466;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Placental calcium-binding protein homolog.

GN S100A4 OR CAPL.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

OC NCBI_TaxID=9913;

OX [1]

RP SEQUENCE.

RC TISSUE=Retina;

RC MEDLINE=94164991; PubMed=8119967;

RA Polans A.S., Palczewski K., Asson-Batres M.A., Ohguro H., Witowska D., Haley T.L., Baizer L., Crab J.W.,

RT "Purification and primary structure of Capl, an S-100-related calcium-binding protein isolated from bovine retina."

RL J. Biol. Chem. 269:6233-6240(1994).

[2]

RP SEQUENCE FROM N.A.

RX MEDLINE=9430466; PubMed=9759666;

RA Duarte W.R., Kaugai S., Iimura T., Oida S., Takenaga K., Ohya K., Ishikawa I.,

RT "cDNA cloning of S100 calcium-binding proteins from bovine periodontal ligament and their expression in oral tissues."

RL J. Dent. Res. 77:1694-1699(1998).

CC -1- SIMILARITY: Belongs to the S-100 family.

CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D89056; BAA13754.1; -

EMBL; A53217; A53217.

HSBP; P30801; 1A03.

InterPro; IPR001751; CAP_S100.

InterPro; IPR002048; EF-hand.

Pfam; PF00036; ehand; 1.

Pfam; PF01023; S_100; 1.

ProDom; PD003407; CAP_S100; 1.

ProDom; PD000012; EF-hand; 1.

ProSite; PS00018; EF_HAND; 1.

ProSite; PS00303; S100 CAP; 1.

KM Calcium-binding; Acetylation.

INIT MET 0

MOD_RES 1

CA_BIND 19 32 ACETYLATION.

CA_BIND 62 73 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).

SEQUENCE 100 AA; 11675 MW; DFCFA7561DBEFA CRC64;

Query Match 13.8%; Score 102.5; DB 1; Length 100;

CC TISSUE=Neutrophils;
 RX MEDLINE=9304974; PubMed=1610833;
 RA Dianoux A.-C.; Stasia M.-J.; Garin J.; Gagnon J.; Vignais P.V.;
 RT "The 23-kilodalton protein, a substrate of protein kinase C, in
 RL bovine neutrophil cytosol is a member of the S100 family.";
 CC Biochemistry 31:5898-5905(1992).
 CC -1- SUBUNIT: Disulfide linked heterodimer of a 7/11 kDa and a 22/23
 CC kDa subunits.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; loosely associated to the
 CC cytoskeleton.
 CC -1- TISSUE SPECIFICITY: Found essentially in phagocytic cells.
 CC -1- PTM: Phosphorylated by protein kinase C.
 CC -1- SIMILARITY: Belongs to the S-100 family.
 CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 DR HSP, P80511, 188A.
 DR InterPro; IPR001751; CaBP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR Prodom; PD003407; CaBP_S100; 1.
 DR Prodom; PD000012; EF-hand; 1.
 DR PROSITE; PS00018; EF_HAND, PARTIAL.
 DR PROSITE; PS00303; S100_CaBP; 1.
 KW Calcium-binding; Phosphorylation.
 FT NON_TER 1 1
 FT CA_BIND 19 32 EF_HAND 1 (LOW AFFINITY) (POTENTIAL).
 FT CA_BIND 63 74 EF_HAND 2 (HIGH AFFINITY) (POTENTIAL).
 SQ SEQUENCE 122 AA, 13673 MW, F3CA8C48806BCCD CRC64;
 Query Match 13.8%; Score 102; DB 1; Length 122;
 Best Local Similarity 31.8%; Pred. No. 0.035;
 Matches 27; Conservative 19; Mismatches 35; Indels 4; Gaps 2;
 QY 54 SLEKAIATTAIPNSSDGKLEKAKD--LLOTFRPAFGQ-ETPKREIISE 109
 DB 2 SGMSEIETITITFQYVRGLGHVDTLQKFKQLVQKELPKKKQKNAINEIMED 61
 QY 110 IDEHTENKIDFEDFMILLSTVMS 134
 DB 62 LDTNVDKQLSFEETLVARLTVAS 86
 RESULT 13
 S10B_BOVIN STANDARD; ERT; 91 AA.
 AC P02638;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE S-100 protein, beta chain.
 GN S100B.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovine; Bos.
 OX NCBI_TaxId=9913;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=79045265; PubMed=710399;
 RA "Isobe T., Okuyama T.,
 RT "The amino-acid sequence of S-100 protein (PAP I-b protein) and its
 RL relation to the calcium-binding proteins.";
 RT Eur. J. Biochem. 89:379-388(1978).
 RN [2]
 RP REVISIONS TO 1.4.
 RX MEDLINE=81236562; PubMed=7250124;
 RA "Isobe T., Okuyama T.,
 RT "The amino-acid sequence of the alpha subunit in bovine brain S-100a
 RL protein.";
 RT Eur. J. Biochem. 116:79-86(1981).
 RN [3]
 RP SEQUENCE.
 RX MEDLINE=85278169; PubMed=4026304;

RA Marshak D.R., Umekawa H., Matterson D.M., Hidaka H.;
 RT "Structural characterization of the calcium binding protein S100 from
 RL adipose tissue.";
 RL Arch. Biochem. Biophys. 240:777-780(1985).
 RP METAL ION-BINDING PROPERTIES.
 RX MEDLINE=84000339; PubMed=6615778;
 RA Buddier J., Gerard D.;
 RT "Ions binding to S100 proteins: structural changes induced by calcium
 RL and zinc on S100a and S100b proteins.";
 RL Biochemistry 22:3360-3369(1983).
 RP CADMIUM-BINDING STUDIES.
 RX MEDLINE=91248136; PubMed=2039467;
 RA Donato H. Jr., Mani R.S., Kay C.M.;
 RT "Spectral studies on the cadmium-ion-binding properties of bovine
 RL brain S-100 protein.";
 RL Biochem. J. 276:13-18(1991).
 RP STRUCTURE BY NMR.
 RX MEDLINE=96398693; PubMed=8805590;
 RA Kilby P.M., van Eudik L.J., Roberts G.C.K.;
 RT "The solution structure of the bovine S100B protein dimer in the
 RL calcium-free state.";
 RL Structure 4:1041-1052(1996).
 CC -1- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 CC distinct binding sites with different affinities exist for both
 CC ions on each monomer. Physiological concentrations of potassium
 CC ion antagonize the binding of both divalent cations, especially
 CC affecting high-affinity calcium-binding sites.
 CC -1- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
 CC one alpha and one beta chain.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Although predominant among the water-soluble
 CC brain proteins, S-100 is also found in a variety of other tissues.
 CC -1- SIMILARITY: Belongs to the S-100 family.
 CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 DR PIR; A91254; BCB01B.
 DR PDB; 1CFP; 12-MAR-97.
 DR PDB; 1MHO; 18-NOV-98.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005576; C:extracellular; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0019240; F:kinase inhibitor activity; NAS.
 DR GO; GO:0042803; F:protein homodimerization activity; ISS.
 DR GO; GO:0048155; F:S100 alpha binding; ISS.
 DR GO; GO:0048154; F:S100 beta binding; ISS.
 DR GO; GO:0048156; F:tau protein binding; IPI.
 DR GO; GO:0008270; F:zinc ion binding; IDA.
 DR GO; GO:0048143; F:astrocyte activation; NAS.
 DR GO; GO:0007409; F:axonogenesis; NAS.
 DR GO; GO:0006749; F:calcium ion homeostasis; ISS.
 DR GO; GO:0006112; F:energy reserve metabolism; ISS.
 DR GO; GO:0048151; F:hyperphosphorylation; NAS.
 DR GO; GO:0006917; P:induction of apoptosis; ISS.
 DR GO; GO:0007611; P:learning and/or memory; ISS.
 DR GO; GO:0045917; P:positive regulation of complement activation; NAS.
 DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
 DR GO; GO:0048159; P:regulation of long-term neuronal synaptic P...; ISS.
 DR GO; GO:0006417; P:regulation of protein biosynthesis; NAS.
 DR InterPro; IPR001751; CaBP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR Prodom; PD003407; CaBP_S100; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CaBP; 1.
 KW Calcium-binding; Zinc; Metal-binding; Acetylation; 3D-structure.
 FT MOD_RES 1 1
 FT CA_BIND 18 31 ACETYLATION.
 FT CA_BIND 61 72 EF_HAND 1 (LOW AFFINITY).
 FT HELIX 2 18 EF_HAND 2 (HIGH AFFINITY).
 FT TURN 19 19

```

FT TURN 24 25
FT STRAND 27 27
FT HELIX 29 39
FT TURN 41 43
FT HELIX 50 60
FT STRAND 68 68
FT HELIX 70 87
SQ SEQUENCE 91 AA; 10537 MW; 386201933DE6B93A CRC64;

Query Match
Best Local Similarity 28.4%; Pred. No. 0.027; DB 1; Length 91;
Matches 23; Conservative 20; Mismatches 35; Indels 3; Gaps 1;

QY 54 SLEKAIATTAIFPNSSDSG--KLEKAIKDLQTFNPAEGQETKREIISLSE 110
DB 1 SLEKAVVALIVFHYSGREGDKKTKSEIKELINNELSHLEIEIKGEVVDKVMETL 60
QY 111 DEHTENKLDPEDEPMILLISIT 131
DB 61 DSDGDECDFOEFMAFAMIT 81
Db

RESULT 14
S104 RAT STANDARD; PRT; 101 AA.
AC P05942;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (Nerve growth factor induced protein 42A) (99K).
GN S100A4
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=88118907; PubMed=3422491;
RA Maslakowski P., Shooter E.M., the genes for two proteins related to a
RT Nerve growth factor induces the genes for two proteins related to a
RT family of calcium-binding proteins in PC12 cells."
RT Proc. Natl. Acad. Sci. U.S.A. 85:1277-1281 (1988).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=88118907; PubMed=3430604;
RA Barricough R., Savin J., Dube S.K., Rudland P.S.,
RT Molecular cloning and sequence of the gene for p9Ka, a cultured
RT myoepithelial cell protein with strong homology to S-100, a calcium-
RT binding protein."
RT J. Mol. Biol. 198:13-20 (1987).
[3]
SEQUENCE OF 3-101 FROM N.A.
RX MEDLINE=92158347; PubMed=1741158;
RA de Vonge M.W., Mukherjee B.B.,
RT Transformation of normal rat kidney cells by v-K-ras enhances
RT expression of transin 2 and an S-100-related calcium-binding
RT protein."
RL Oncogene 7:109-119 (1992).
CC -1- INDUCTION: By nerve growth factor.
CC -1- SIMILARITY: Belongs to the S-100 family.
CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X06916; CAA30014.1; -
CC EMBL; J03628; AAA42098.1; -

```

```

DR EMBL; X64022; -; NOT ANNOTATED_CDS.
DR EMBL; X64023; -; NOT ANNOTATED_CDS.
DR PIR; S01759; S01759.
DR HSSP; P30801; 1A03.
DR InterPro; IPR001751; CAP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand.1.
DR ProDom; PD003407; CAP_S100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CAP; 1.
KM Calcium-binding.
FT CA BIND 63 74
FT CA BIND 20 33
SQ SEQUENCE 101 AA; 11776 MW; EA0619CEB4F487C1 CRC64;

Query Match
Best Local Similarity 25.9%; Pred. No. 0.046; DB 1; Length 101;
Matches 21; Conservative 24; Mismatches 33; Indels 3; Gaps 1;

QY 56 LEKAIATTAIFPNSSDSG--KLEKAIKDLQTFNPAEGQETKREIISLSE 112
DB 5 LEKALDVIVSTFHRYSGNEGDKPKNTKELITRELPSPFLGRTRDEAFAQKMMNLD 64
QY 113 HTENKLDPEDEPMILLISITVM 133
DB 65 NRDEVDFOERYCVFLCIAVM 85
Db

RESULT 15
S10B HUMAN STANDARD; PRT; 91 AA.
AC P04271;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S-100 protein, beta chain.
GN S100B
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
[1]
SEQUENCE.
RX MEDLINE=85291729; PubMed=4031854;
RA Jensen R., Marshak D.R., Anderson C., Lukas T.J., Watterson D.M.;
RT "Characterization of human brain S100 protein fraction: amino acid
RT sequence of S100 beta."
RL J. Neurochem. 45:700-705 (1985).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=90368757; PubMed=2394738;
RA Allore R.J., Friend W.C., O'Hanlon D., Neilson K.M., Bauml R.,
RA Dunn R.J., Marks A.;
RT "Cloning and expression of the human S100 beta gene."
RL J. Biol. Chem. 265:15537-15543 (1990).
[3]
SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Schat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,
RA Ditchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong D.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schietz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Cammici P., Prange C.,
RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:21:57 ; Search time 44 Seconds
(without alignments)
939.927 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 741

Sequence: 1 MCGCNHSTQDKSLHLEGPEN.....LSITVMSDLLQINRVKIMK 147

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US03_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	690	93.1	140	9	US-09-764-864-1447
2	131.5	17.7	94	14	US-10-097-340-270
3	123	16.6	97	14	US-10-097-340-274
4	123	16.6	97	14	US-10-171-311-206
5	123	16.6	97	15	US-10-236-031B-40
6	108.5	14.6	134	10	US-09-992-600A-2
7	108.5	14.6	134	10	US-09-924-340-2
8	108.5	14.6	134	10	US-09-992-095B-2
9	108.5	14.6	134	10	US-09-999-570-2
10	108.5	14.6	134	14	US-10-000-489-2
11	108.5	14.6	134	14	US-10-000-986-2
12	108.5	14.6	134	14	US-10-154-678-2
13	106.5	14.4	101	9	US-09-393-433-1
14	106.5	14.4	101	9	US-09-781-509-1
15	106.5	14.4	101	12	US-10-087-192-1158

16	106.5	14.4	101	13	US-10-067-618-2	Sequence 2, Appli
17	106.5	14.4	101	13	US-10-135-152-2	Sequence 2, Appli
18	106.5	14.4	101	14	US-10-269-643-1	Sequence 1, Appli
19	106.5	14.4	134	9	US-09-925-302-694	Sequence 694, App
20	106.5	14.4	134	12	US-09-325-302-694	Sequence 694, App
21	105.5	14.2	119	12	US-10-087-192-1155	Sequence 1155, Ap
22	103	13.9	186	15	US-10-264-049-4104	Sequence 4104, Ap
23	102.5	13.8	101	9	US-09-393-433-2	Sequence 2, Appli
24	102.5	13.8	101	9	US-09-781-509-2	Sequence 2, Appli
25	102.5	13.8	101	14	US-10-269-643-2	Sequence 2, Appli
26	100	13.5	118	15	US-10-264-049-3289	Sequence 3289, Ap
27	98.5	13.3	92	10	US-09-492-026-5	Sequence 5, Appli
28	98.5	13.3	92	10	US-09-319-039-184	Sequence 184, App
29	98.5	13.3	92	12	US-10-336-603A-102	Sequence 102, App
30	96	13.0	95	10	US-09-877-843-27	Sequence 27, Appl
31	96	13.0	96	10	US-09-877-843-25	Sequence 25, Appl
32	96	13.0	97	11	US-09-877-843-26	Sequence 26, Appl
33	96	13.0	97	11	US-09-997-003-32	Sequence 32, Appl
34	96	13.0	97	14	US-10-304-287-3	Sequence 3, Appli
35	96	13.0	112	14	US-10-305-219-161	Sequence 161, App
36	96	13.0	113	10	US-09-492-026-7	Sequence 7, Appli
37	95	12.8	97	10	US-09-877-843-29	Sequence 29, Appl
38	91.5	12.3	89	14	US-10-134-841-1	Sequence 1, Appli
39	91.5	12.3	218	12	US-10-296-115-772	Sequence 772, App
40	91	12.3	97	10	US-09-877-843-28	Sequence 28, Appl
41	90.5	12.2	89	14	US-10-316-253-46	Sequence 46, Appl
42	89.5	12.1	95	9	US-09-919-172-102	Sequence 102, App
43	89.5	12.1	95	9	US-09-981-353-98	Sequence 98, Appl
44	89.5	12.1	113	12	US-10-276-774-2377	Sequence 2377, Ap
45	89	12.0	93	15	US-10-094-886-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-09-764-864-1447
; Sequence 1447, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1447
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1447

Query Match 93.1%; Score 690; DB 9; Length 140;
Best Local Similarity 99.3%; Pred. No. 1.4e-64;
Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	9	QDKSLHLEGDNPSPAPTCTCAPKMPKRSISIKOLASVKALRKCSDLEKAIATTAIFR	68
DB	2	EDKSLHLEGDNPSPAPTCTCAPKMPKRSISIKOLASVKALRKCSDLEKAIATTAIFR	61
QY	69	NSSDSGKLEKAIKDLLQTOFRNFAEQETPKYRIILSELDEHTENKLDFFEDFMILL	128
DB	62	NSSDSGKLEKAIKDLLQTOFRNFAEQETPKYRIILSELDEHTENKLDFFEDFMILL	121
QY	129	SITVMSDLLQINRVKIMK	147
DB	122	SITVMSDLLQINRVKIMK	140

RESULT 2
US-10-097-340-270

```

/ APPLICANT: Peter VEIBY
/ APPLICANT: Gordon B. MILLS
/ APPLICANT: Robert C. BAST, Jr.
/ APPLICANT: Karen LU
/ APPLICANT: Rosemarie SCHMANDT
/ APPLICANT: Xumei ZHAO
/ APPLICANT: Karen GLATT
/ TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
/ TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
/ FILE REFERENCE: MRI-030
/ CURRENT APPLICATION NUMBER: US/10/097,340
/ CURRENT FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 60/276,025
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/325,149
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 60/276,026
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/324,967
/ PRIOR FILING DATE: 2001/09/26
/ PRIOR APPLICATION NUMBER: 60/311,732
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/325,102
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 60/323,580
/ PRIOR FILING DATE: 2001-09-19
/ NUMBER OF SEQ ID NOS: 363
/ SOFTWARE: Fast-Seq for Windows Version 4.0
/ SEQ ID NO 274
/ LENGTH: 97
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-097-340-274

Query Match 16.6%; Score 123; DB 14; Length 97;
Best Local Similarity 31.1%; P2; Mismatches 36; Indels 4; Gaps 2;
Matches 28; Conservative

QY 53 CSDLKALATATLIFRNSDSBQ---KLEKALAKDLLQTFRNFAGQETPKPYRELLSE 109
DB 2 CSSLEQALAVLVITTFHKYSQCSGDKFKLSKGMKELLHKELPSPFVGKVDGGLKLMGS 61
QY 110 LDEHTENKLDPEDFMILLISITVM-SDLLQ 138
DB 62 LDENSQQQVDFQYAVFLALITWNCNDFPQ 91

RESULT 4
US-10-171-311-206
/ Sequence 206, Application US/10171311
/ Publication No. US20030087270A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Chen, Yan
/ APPLICANT: Zhao, Xumei
/ APPLICANT: Monahan, John
/ APPLICANT: Kamatkar, Shubhangi
/ APPLICANT: Glatt, Karen
/ APPLICANT: Gamavaraapu, Manjula
/ APPLICANT: Hoersd, Sebastian
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
/ TITLE OF INVENTION: OF CERVICAL CANCER
/ FILE REFERENCE: MRI-035
/ CURRENT APPLICATION NUMBER: US/10/171,311
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: US 60/298,159
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,155
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/335,936
/ PRIOR FILING DATE: 2001-11-14
/ NUMBER OF SEQ ID NOS: 238

```

Fri Jun 4 16:16:01 2004

```
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 206
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-206

Query Match      16.6%; Score 123; DB 14; Length 97;
Best Local Similarity 31.1%; Pred. No. 4.1e-05;
Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;

Qy 53 CSDELEKAIATATLIFRNSSDG---KLEKAIADLLLOTFRNFAEQETPKYREILSE 109
Db 2 CSSLEQALAVLVTFHKYSCQEGDKFKLSGEMKELLHKELPSFVGKVDDEGLKLMGS 61

Qy 110 LDEHTENKLDPEDFMILLLSITVM-SDLLQ 138
Db 62 LDENSQQQVDFQYAVFLALITVMCNDFQ 91

RESULT 5
US-10-236-031B-40
; Sequence 40, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-40

Query Match      16.6%; Score 123; DB 15; Length 97;
Best Local Similarity 31.1%; Pred. No. 4.1e-05;
Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;

Qy 53 CSDELEKAIATATLIFRNSSDG---KLEKAIADLLLOTFRNFAEQETPKYREILSE 109
Db 2 CSSLEQALAVLVTFHKYSCQEGDKFKLSGEMKELLHKELPSFVGKVDDEGLKLMGS 61

Qy 110 LDEHTENKLDPEDFMILLLSITVM-SDLLQ 138
Db 62 LDENSQQQVDFQYAVFLALITVMCNDFQ 91

RESULT 6
US-09-992-600A-2
; Sequence 2, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06

; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 206
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-206

Query Match      16.6%; Score 123; DB 14; Length 97;
Best Local Similarity 31.1%; Pred. No. 4.1e-05;
Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;

Qy 53 CSDELEKAIATATLIFRNSSDG---KLEKAIADLLLOTFRNFAEQETPKYREILSE 109
Db 2 CSSLEQALAVLVTFHKYSCQEGDKFKLSGEMKELLHKELPSFVGKVDDEGLKLMGS 61

Qy 110 LDEHTENKLDPEDFMILLLSITVM-SDLLQ 138
Db 62 LDENSQQQVDFQYAVFLALITVMCNDFQ 91

RESULT 7
US-09-924-340-2
; Sequence 2, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 2
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-2

Query Match      14.6%; Score 108.5; DB 10; Length 134;
Best Local Similarity 25.2%; Pred. No. 0.0022;
Matches 27; Conservative 23; Mismatches 52; Indels 5; Gaps 1;

Qy 26 TSTCAPRKMPKRISISKQLASVKALRKCSDELEKAIATATLIFRNSSDGKLEKAIADL 85
Db 17 TFSCLPRPRTEPL-----VASTDHTKMPSQMEHAMETWMTFHKFAGDKGYLTKEIDLRLV 71

Qy 86 LOTQFRNFAEQETPKYREILSELDHTENKLDPEDFMILLLSITV 132
Db 72 MEKEFFPGFLENQKDPPLAVDKIMKDLQCRDQKGVQSFPSLIAGLTI 118

RESULT 8
US-09-992-095B-2
; Sequence 2, Application US/09992095B
; Publication No. US20030157485A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US5.DIV
; CURRENT APPLICATION NUMBER: US/09/992,095B
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 134
; ORGANISM: Homo sapiens
; US-09-992-095B-2

```

```

Query Match 14.6%; Score 108.5; DB 10; Length 134;
Best Local Similarity 25.2%; Pred. No. 0.0022;
Matches 27; Conservative 23; Mismatches 52; Indels 5; Gaps 1;

Qy 26 TSTCAPKMPKRISSIKQLASVKALRKCSDLKAIATTAIFRNSSDSGKLEKAIADL 85
Db 17 TFSCLPRPRTEPL-----VASTDHTKMPQMEHMETMTFTFKFAGDKGYLTKEIDLRLV 71

Qy 86 LQTOFRNFAEGQETPKYREILSELDEHTENKLDPEDFMILLISITV 132
Db 72 MEKEFFPGFLENQKDLAVDKIMKLDQCRDGKVGQSFPSLIAGLTI 118

```

```

RESULT 9
US-09-999-570-2
; Sequence 2, Application US/09999570
; Publication No. US20030170628A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: G-091US08DIV
; CURRENT APPLICATION NUMBER: US/09/999,570
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 2
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-999-570-2

```

```

Query Match 14.6%; Score 108.5; DB 10; Length 134;
Best Local Similarity 25.2%; Pred. No. 0.0022;

```

```

Matches 27; Conservative 23; Mismatches 52; Indels 5; Gaps 1;

Qy 26 TSTCAPKMPKRISSIKQLASVKALRKCSDLKAIATTAIFRNSSDSGKLEKAIADL 85
Db 17 TFSCLPRPRTEPL-----VASTDHTKMPQMEHMETMTFTFKFAGDKGYLTKEIDLRLV 71

Qy 86 LQTOFRNFAEGQETPKYREILSELDEHTENKLDPEDFMILLISITV 132
Db 72 MEKEFFPGFLENQKDLAVDKIMKLDQCRDGKVGQSFPSLIAGLTI 118

```

```

RESULT 10
US-10-000-489-2
; Sequence 2, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 2
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-000-489-2

```

```

Query Match 14.6%; Score 108.5; DB 14; Length 134;
Best Local Similarity 25.2%; Pred. No. 0.0022;
Matches 27; Conservative 23; Mismatches 52; Indels 5; Gaps 1;

Qy 26 TSTCAPKMPKRISSIKQLASVKALRKCSDLKAIATTAIFRNSSDSGKLEKAIADL 85
Db 17 TFSCLPRPRTEPL-----VASTDHTKMPQMEHMETMTFTFKFAGDKGYLTKEIDLRLV 71

Qy 86 LQTOFRNFAEGQETPKYREILSELDEHTENKLDPEDFMILLISITV 132
Db 72 MEKEFFPGFLENQKDLAVDKIMKLDQCRDGKVGQSFPSLIAGLTI 118

```

```

RESULT 11
US-10-000-986-2
; Sequence 2, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US/10/000,986
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277

```

; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/299,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 2
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-000-986-2

Query Match 14.6%; Score 108.5; DB 14; Length 134;
Best Local Similarity 25.2%; Pred. No. 0.0022;
Matches 27; Conservative 23; Mismatches 52; Indels 5; Gaps 1;
Qy 26 TSTCAPRMPKRISIKQLASVKALRKCSDLEKAIATATLIFRNSSDSGKLEKAIKOL 85
Db 17 TFSCLPRPRTEPL-----VASTDHTKMPQSQMEHAMETMTFTFKFAGDKGLTKEDLRVL 71
Qy 86 LQTOFRNFAEGQETPKPKYREILSELDEHTENKLDDEDPMILLISITV 132
Db 72 MEKEFPGFLENQKDPFLAVDKIMKOLDQCRDQKGVGFQSFSLIAGLTI 118

RESULT 12

US-10-154-678-2
; Sequence 2, Application US/10154678
; Publication No. US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.BEG
; CURRENT APPLICATION NUMBER: US/10/154,678
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 2
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-154-678-2

Query Match 14.6%; Score 108.5; DB 14; Length 134;
Best Local Similarity 25.2%; Pred. No. 0.0022;
Matches 27; Conservative 23; Mismatches 52; Indels 5; Gaps 1;
Qy 26 TSTCAPRMPKRISIKQLASVKALRKCSDLEKAIATATLIFRNSSDSGKLEKAIKOL 85
Db 17 TFSCLPRPRTEPL-----VASTDHTKMPQSQMEHAMETMTFTFKFAGDKGLTKEDLRVL 71
Qy 86 LQTOFRNFAEGQETPKPKYREILSELDEHTENKLDDEDPMILLISITV 132
Db 72 MEKEFPGFLENQKDPFLAVDKIMKOLDQCRDQKGVGFQSFSLIAGLTI 118

RESULT 13

US-09-393-433-1
; Sequence 1, Application US/09393433
; Patent No. US2001001128A1
; GENERAL INFORMATION:

; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/393,433
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-393-433-1

Query Match 14.4%; Score 106.5; DB 9; Length 101;
Best Local Similarity 28.4%; Pred. No. 0.0024;
Matches 23; Conservative 22; Mismatches 33; Indels 3; Gaps 1;
Qy 56 LEKAIATATLIFRNSSDSG---KLEKAIKOLLQTOFRNFAEGQETPKPKYREILSELDE 112
Db 5 LEKALDVMVSTFHKYSGKEGDKFKLKSSELKELLTRELPSFLGKRTDEAAFOKLMSNLD 64
Qy 113 HTENKLDDEDPMILLISITVM 133
Db 65 NRDNEVDFOEYCVFLSIAMM 85

RESULT 14

US-09-781-509-1
; Sequence 1, Application US/09781509
; Patent No. US20020099010A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-781-509-1

Query Match 14.4%; Score 106.5; DB 9; Length 101;
Best Local Similarity 28.4%; Pred. No. 0.0024;
Matches 23; Conservative 22; Mismatches 33; Indels 3; Gaps 1;
Qy 56 LEKAIATATLIFRNSSDSG---KLEKAIKOLLQTOFRNFAEGQETPKPKYREILSELDE 112
Db 5 LEKALDVMVSTFHKYSGKEGDKFKLKSSELKELLTRELPSFLGKRTDEAAFOKLMSNLD 64
Qy 113 HTENKLDDEDPMILLISITVM 133
Db 65 NRDNEVDFOEYCVFLSIAMM 85

RESULT 15

US-10-087-192-1158
; Sequence 1158, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122


```

; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1158
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1158

```

```

Query Match      14.4%; Score 106.5; DB 12; Length 101;
Best Local Similarity 28.4%; Pred. No. 0.0024;
Matches 23; Conservative 22; Mismatches 33; Indels 3; Gaps 1;

QY      56 LEKATATTALIFRNSSDSG--KLEKAIADLLQTFRPAEQGQETPKPYREILSELDE 112
Db      5 LEKALDVWVSTPHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAFAQKLSNLD 64

QY      113 HTENKLDPEDFMILLISITVM 133
Db      65 NRDNEVDVFOEYCVFLSCIAMM 85

```

Search completed: June 3, 2004, 13:27:49
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:19:26 ; Search time 22 Seconds
(without alignments)
344.955 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 741

Sequence: 1 MGGCMHSTQDKSLHLEDPN.....LSITVMSDLLQIRNVKIMK 147

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/PTCTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfileal.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	741	100.0	147	4	US-09-621-976-7168
2	123	16.6	97	1	US-07-662-198B-2
3	106.5	14.4	101	1	US-08-190-560-2
4	106.5	14.4	101	1	US-08-469-277-2
5	106.5	14.4	101	2	US-08-468-946-2
6	106.5	14.4	101	2	US-08-468-942-2
7	106.5	14.4	101	4	US-09-298-625-2
8	102.5	13.8	92	2	US-09-051-589-1
9	98.5	13.3	91	1	US-07-987-272A-11
10	98.5	13.3	92	2	US-08-918-727-5
11	98.5	13.3	92	3	US-09-205-680A-5
12	96	13.0	95	1	US-07-987-272A-5
13	96	13.0	113	2	US-08-918-727-7
14	96	13.0	113	3	US-09-205-680A-7
15	91.5	12.3	88	1	US-07-987-272A-14
16	91.5	12.3	89	1	US-07-987-272A-12
17	89.5	12.1	95	4	US-09-919-172-102
18	89.5	12.1	95	4	US-09-976-594-467
19	89	12.0	95	4	US-09-399-913-65
20	89	12.0	114	1	US-08-385-241-3
21	89	12.0	114	4	US-09-214-272-4
22	86	11.6	75	1	US-07-987-272A-17
23	83.5	11.3	76	1	US-07-987-272A-12
24	79	10.7	109	1	US-07-987-272A-8
25	77	10.4	1664	1	US-09-599-652-2
26	77	10.4	1664	2	US-08-642-846-2
27	77	10.4	1664	4	US-09-264-604-2

28	76.5	10.3	105	2	US-08-918-727-6	Sequence 6, Appli
29	76.5	10.3	105	3	US-09-205-680A-6	Sequence 6, Appli
30	76	10.3	391	3	US-08-773-731A-5	Sequence 5, Appli
31	74	10.0	196	4	US-09-621-976-5041	Sequence 5041, Ap
32	74	10.0	222	4	US-09-205-258-360	Sequence 360, App
33	74	10.0	311	4	US-09-252-991A-23527	Sequence 23527, A
34	74	10.0	456	1	US-08-464-164-2	Sequence 2, Appli
35	74	10.0	456	1	US-08-338-057-2	Sequence 2, Appli
36	74	10.0	456	1	US-08-668-416-2	Sequence 2, Appli
37	73	9.9	328	4	US-09-252-991A-27822	Sequence 27822, A
38	73	9.9	521	2	US-08-406-855A-19	Sequence 19, Appl
39	73	9.9	521	3	US-09-206-899-19	Sequence 19, Appl
40	73	9.9	699	4	US-09-457-040B-18	Sequence 18, Appl
41	72.5	9.8	1093	3	US-09-315-793-52	Sequence 52, Appl
42	72	9.7	101	1	US-08-469-486-58	Sequence 58, Appl
43	72	9.7	101	2	US-08-469-658-58	Sequence 58, Appl
44	72	9.7	230	4	US-09-252-991A-26110	Sequence 26110, A
45	71.5	9.6	1394	4	US-09-845-917A-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-7168
; Sequence 7168, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7168
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7168

Query Match 100.0%; Score 741; DB 4; Length 147;
Best Local Similarity 100.0%; Pred No. 7,6e-73;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGGCMHSTQDKSLHLEDPNPSAAPTSTCAPRMPKRISISKOLASVKALRKCSDEKAI	60
Db	1	MGGCMHSTQDKSLHLEDPNPSAAPTSTCAPRMPKRISISKOLASVKALRKCSDEKAI	60
Qy	61	ATTALIFRNSSDGKLEKAIKDLLOTFRPAEQETPKPYREILSLDHTENKLDIF	120
Db	61	ATTALIFRNSSDGKLEKAIKDLLOTFRPAEQETPKPYREILSLDHTENKLDIF	120
Qy	121	EDFMILLISITVMSDLLQIRNVKIMK	147
Db	121	EDFMILLISITVMSDLLQIRNVKIMK	147

RESULT 2

US-07-662-198B-2
; Sequence 2, Application US/07662198B
; Patent No. 5262528
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Lee, Sam W.
; APPLICANT: Tomasetto, Catherine
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street

```
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,198B
; FILING DATE: 19910228
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/049001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; US-07-662-198B-2

Query Match 16.6%; Score 123; DB 1; Length 97;
Best Local Similarity 31.1%; Pred. No. 5.5e-06;
Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;

QY 53 CSDEKAIATTAIFRNSSDGDG---KLEKAIARDLLQTOFRNFAEGQETPKPKYRIELSE 109
Db 2 CSSLEQALAVITTFHKYSCQEGDKFKLSKGENKELLHKELPSFVGEKVDEGLKLMGN 61
QY 110 LDEHTENKLPEDFMILLISITVM-SLLQ 138
Db 62 LDENSQQQVDFQYAVFLALITVNCNDFQ 91

RESULT 3
US-08-130-560-2
; Sequence 2, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-130-560-2

; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,198B
; FILING DATE: 19910228
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/049001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; US-07-662-198B-2

Query Match 14.4%; Score 106.5; DB 1; Length 101;
Best Local Similarity 28.4%; Pred. No. 0.00036;
Matches 23; Conservative 22; Mismatches 33; Indels 3; Gaps 1;

QY 56 LEKAIATTAIFRNSSDGDG---KLEKAIARDLLQTOFRNFAEGQETPKPKYRIELSE 112
Db 5 LEKALDVNVSTFHKYSGEGDKFKLSKSELKELLTRELPSFLGKRTDEAAAFQKLMSNLS 64
QY 113 HTENKLPEDFMILLISITVM 133
Db 65 NRDNVDVDFQYCVFLSCIAMM 85

RESULT 4
US-08-469-277-2
; Sequence 2, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,277
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-277-2
```

RESULT 6
US-08-468-942-2
; Sequence 2, Application US/08468942
; Patent No. 5965360
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530

STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,942
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PS-08-468-942-2

RESULT 7
US-09-298-625-2
; Sequence 2, Application US/09298625
; Patent No. 6638504
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YXII-2
; CURRENT APPLICATION NUMBER: US/09/298,625
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31

; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-625-2

Query Match 14.4%; Score 106.5; DB 4; Length 101;
Best Local Similarity 28.4%; Pred. No. 0.00036;
Matches 23; Conservative 22; Mismatches 33; Indels 3; Gaps 1;

QY 56 LEKAIATATTALIFRNSSDSG---KLEKAIADLLQTFQFNFAEGQETPKPKYREILSELDE 112
DB 5 LEKALDVMVSTFKYSGKGDGFKLNKSELKELLIRELPSFLGKRTDEAFAFKLMSNLDS 64

QY 113 HTENKLDFFEDFMILLISITVM 133
DB 65 NRDNVEDFQBYCVFLSCIAMM 85

RESULT 8

US-09-051-589-1
; Sequence 1, Application US/09051589
; Patent No. 5990080
; GENERAL INFORMATION:

; APPLICANT: HAGLID, Kenneth G.
; TITLE OF INVENTION: USE OF PROTEIN S-100B IN MEDICINES CONTAINING THE
; TITLE OF INVENTION: PROTEIN S-100B
; FILE REFERENCE: 003300-478

; CURRENT APPLICATION NUMBER: US/09/051,589
; CURRENT FILING DATE: 1998-04-15
; EARLIER APPLICATION NUMBER: SE 9503620-8
; EARLIER FILING DATE: 1995-10-17
; EARLIER APPLICATION NUMBER: PCT/SE96/01305
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Protein S-100b
US-09-051-589-1

Query Match 13.8%; Score 102.5; DB 2; Length 92;
Best Local Similarity 28.4%; Pred. No. 0.00086;
Matches 23; Conservative 21; Mismatches 34; Indels 3; Gaps 1;

QY 54 SDLEKAIATATTALIFRNSSDSG---KLEKAIADLLQTFQFNFAEGQETPKPKYREILSEL 110
DB 2 SELEKAWALIDVFHQYSGREGDKHKKSELKELINNELSHFLEBIKEQEVVDKWEITL 61

QY 111 DEHTENKLDFFEDFMILLISIT 131
DB 62 DEDGDGECDFQEFMAFVSWT 82

RESULT 9

US-07-987-272A-11
; Sequence 11, Application US/07987272A
; Patent No. 5731166
; GENERAL INFORMATION:

; APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman Darby & Cushman
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
; CITY: Washington

; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/987,272A
; FILING DATE: 05-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 2127
; FILING DATE: 05-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 4463
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brinkman, David W
; REGISTRATION NUMBER: 20,817
; REFERENCE/DOCKET NUMBER: DWB/1925/200259
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861 3000
; TELEFAX: 202-822 0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-987-272A-11

Query Match 13.3%; Score 98.5; DB 1; Length 91;
Best Local Similarity 27.2%; Pred. No. 0.0023;
Matches 22; Conservative 21; Mismatches 35; Indels 3; Gaps 1;

QY 54 SDLEKAIATATTALIFRNSSDSG---KLEKAIADLLQTFQFNFAEGQETPKPKYREILSEL 110
DB 1 SELEKAWALIDVFHQYSGREGDKHKKSELKELINNELSHFLEBIKEQEVVDKWEITL 60

QY 111 DEHTENKLDFFEDFMILLISIT 131
DB 61 DNDGDGECDFQEFMAFVAMVT 81

RESULT 10

US-08-918-727-5
; Sequence 5, Application US/08918727
; Patent No. 5849528
; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN S100 PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

; REFERENCE/DOCNET NUMBER: PF-0373 US

US-07-987-272A-9

Query Match 13.0%; Score 96; DB 1; Length 95;
Best Local Similarity 26.6%; Pred. No. 0.0045;
Matches 21; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

QY 54 SLEKAIATTAIFRNSSDSGKLEKAIADLLQTFRNFAEGQETPKYREILSELDEH 113
DB 2 SOVEHAMEMTTFKFKACKGYLKEDLRLVMEKEFPGLFNQKDLPLANDKIMKLDQC 61

QY 114 TENKLDFFEDFMILLISITV 132
DB 62 RDCKVGQSFPSLIAGLTI 80

RESULT 13

US-08-918-727-7
; Sequence 7, Application US/08918727
; Patent No. 5849528
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN S100 PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/918,727
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0373 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 488157

Query Match 13.0%; Score 96; DB 2; Length 113;
Best Local Similarity 30.1%; Pred. No. 0.0059;
Matches 25; Conservative 20; Mismatches 34; Indels 4; Gaps 2;

QY 49 ALRKCSDEKAIATTAIFRNSSDSG---KLEKAIADLLQTFRNFAEGQETPKYR 104
DB 2 AAKTGSQERSISTINVFHYSRKYGHPTLNKAEFKEMVKNKDLNFKREKRNENLLR 61

QY 105 EILSELDEHTENKLDFFEDFMILL 127
DB 62 DIMEDLTNQDNQLSFECCWMLM 84

RESULT 14

US-09-205-680A-7
; Sequence 7, Application US/09205680A
; Patent No. 6103497
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN S100 PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/205,680A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Colette C. Muenzen
; REGISTRATION NUMBER: 39,784
; REFERENCE/DOCKET NUMBER: PF-0373 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 486157

Query Match 13.0%; Score 96; DB 3; Length 113;
Best Local Similarity 30.1%; Pred. No. 0.0058;
Matches 25; Conservative 20; Mismatches 34; Indels 4; Gaps 2;

QY 49 ALRKCSDEKAIATTAIFRNSSDSG---KLEKAIADLLQTFRNFAEGQETPKYR 104
DB 2 AAKTGSQERSISTINVFHYSRKYGHPTLNKAEFKEMVKNKDLNFKREKRNENLLR 61

QY 105 EILSELDEHTENKLDFFEDFMILL 127
DB 62 DIMEDLTNQDNQLSFECCWMLM 84

RESULT 15

US-07-987-272A-1
; Sequence 1, Application US/07987272A
; Patent No. 5731166
; GENERAL INFORMATION:

Fri Jun 4 16:16:01 2004

APPLICANT: Gezy, C., Simpson, R. J. and Lackmann, M
TITLE OF INVENTION: No. 5731166el Chemotactic Factor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman Darby & Cushman
STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/987,272A
FILING DATE: 05-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK 2127
FILING DATE: 05-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK 4463
FILING DATE: 05-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Brinkman, David W
REGISTRATION NUMBER: 20,817
REFERENCE/DOCKET NUMBER: DWB/1925/200259
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861 3000
TELEFAX: 202-822 0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 88 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-987-272A-1

Query Match 12.3%; Score 91.5; DB 1; Length 88;
Best Local Similarity 27.4%; Pred. No. 0.013;
Matches 23; Conservative 23; Mismatches 31; Indels 7; Gaps 2;
QY 54 SDLEKAIATATLIFNSSDSQK--LEKAIKDLLOTPRNFAGQETPKYREILSEL 110
DB 2 SELEKALNLDIVHYNISIQNHLYNDFKKAVTTECPQVQNIENLFR----EL 57
QY 111 DEHTENKLDFFEDMILLSITVMS 134
DB 58 DINSNAINFEFLAWIKVGVAS 81

Search completed: June 3, 2004, 13:23:02
Job time : 23 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:14:01 ; Search time 56 seconds
(without alignments)
741.687 Million cell updates/sec

Title: US-09-744-197-1
Perfect score: 741
Sequence: 1 MG3CMHSTQKSLHLEDPN.....LSITVMSDLLQNRNWKIMK 147

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	741	100.0	147	3	AA58793 Human cal
2	715	96.5	143	4	AA40026 Human pol
3	690	93.1	140	4	AAU16494 Human nov
4	690	93.1	140	6	ABU55563 Human nov
5	686	92.6	137	4	AA41812 Human pol
6	138.5	18.7	93	7	ABE61762 Rat Prote
7	131.5	17.7	93	7	ABE61764 Human Pro
8	131.5	17.7	94	3	AA45531 Human S10
9	131.5	17.7	94	4	AA40258 Human pol
10	131.5	17.7	94	5	ABG96406 Human ova
11	123	16.6	97	2	AA27058 Sequence
12	123	16.6	97	2	AA26406 Sequence
13	123	16.6	97	3	AA45532 Human S10
14	123	16.6	97	5	ABG96408 Human ova
15	123	16.6	97	6	ABR92148 Human cer
16	123	16.6	97	7	ABR70348 S100 calc
17	123	16.6	97	7	AD59560 Human Pro
18	123	16.6	97	7	AD59556 Human Pro
19	123	16.6	98	6	ABP71986 Human sta
20	123	16.6	98	6	ABU56414 Lung can
21	123	16.6	98	6	ABU56415 Lung can
22	123	16.6	98	6	ABU56412 Lung can
23	111.5	15.0	124	7	ABR08957 Novel pro
24	108.5	14.6	134	6	ABR48453 Human S-1
25	106.5	14.4	101	2	AA20560 Human mts

26	106.5	14.4	101	2	AA80453 Human mts
27	106.5	14.4	101	3	AA45534 Human S10
28	106.5	14.4	101	3	AA37432 Human mts
29	106.5	14.4	101	4	AA72386 Human mts
30	106.5	14.4	101	6	ABU08513 Human mts
31	106.5	14.4	101	7	ADD14157 Human src
32	106.5	14.4	134	3	AA58356 Lung can
33	103	13.9	186	5	ABP42972 Human ova
34	102.5	13.8	97	5	ABP51390 Human MDD
35	102.5	13.8	97	6	ABU11560 Human MDD
36	102.5	13.8	101	4	AB72387 Murine Mc
37	100	13.5	118	5	ABP42157 Human ova
38	98.5	13.3	92	5	ABR97495 Novel hum
39	98.5	13.3	92	6	ABR58703 Human can
40	98.5	13.3	92	6	ABU63333 Human S10
41	98.5	13.3	92	7	ADD49129 Human NOV
42	98.5	13.3	92	8	ADE77019 Human pro
43	98.5	13.3	153	7	ADE07872 Novel pro
44	98	13.2	152	4	AA39994 Human pol
45	96.5	13.0	91	2	AA46607 Human bra

ALIGNMENTS

RESULT 1

AA58793
ID AA58793 standard; protein; 147 AA.

XX
AC AA58793;

DT
DT 08-MAY-2000 (first entry)

XX
DE Human calcium regulatory protein CaREG-1.

XX
KW Human; calcium regulatory protein 1; CaREG-1; anticonvulsant; vasotropic; cerbroprotective; neurotropic; neuroprotective; neuroleptic; cardiant; cyotatic; epilepsy; stroke; Alzheimer's disease; cardiomyopathy; myocarditis; Duchenne's muscular dystrophy; tachyarrhythmia; cancer; diagnosis; therapy.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT Region 1..98 /note= "similarity to annexin type I signature"

FT Modified-site 7 /note= "potential O-phosphorylation"

FT Modified-site 39 /note= "potential O-phosphorylation"

FT Modified-site 54 /note= "potential O-phosphorylation"

FT Region 60..144 /note= "similarity to S100 Ca2+ binding protein signature"

FT Domain 60..144 /note= "similarity to EF-hand calcium binding domain"

FT Modified-site 69 /note= "potential N-glycosylation"

FT Modified-site 71 /note= "potential O-phosphorylation"

FT Modified-site 108 /note= "potential O-phosphorylation"

FT Region 115..128 /note= "similarity to annexin type I signature"

XX
WO200005368-A1.

XX
PD 03-FEB-2000.

XX
EF 19-JUL-1999; 99WO-US012385.

XX
PR 20-JUL-1998; 98US-0160076P.

XX PA (INCY-) INCYTE PHARM INC.
XX PI Bandman O, Tang YT, Corley NC, Guegler KJ, Baughn MR, Junming Y;
XX PR MPI; 2000-182695/16.
XX DR N-PSDB; AAZ58233.
XX PT New regulatory proteins and polynucleotides useful for treating and
XX PT detecting neurological disorders including epilepsy, Alzheimer's disease,
XX PT and stroke and cardiovascular disorders e.g. cardiomyopathy, myocarditis,
XX PT tachyarrhythmia.
XX PS Claim 1; Fig 1A-B; 74pp; English.
XX CC The present sequence is that of novel human calcium regulatory protein 1
XX CC (CAREG-1), as deduced from a consensus sequence (see AAZ58233) of
XX CC isolated cDNA clones. The protein has chemical and structural similarity
XX CC with human S100-alpha. Northern analysis showed that CAREG-1 is expressed
XX CC exclusively in lung tissue, including foetal and cancerous lung tissue.
XX CC The invention provides CAREG-1 and CAREG-2 polynucleotides and
XX CC polypeptides, expression vectors, host cells, antibodies, agonists and
XX CC antagonists. The polypeptides, polynucleotides and agonists are useful
XX CC for treating or preventing neurological disorders such as epilepsy,
XX CC ischaemic cerebrovascular diseases, stroke, Alzheimer's disease etc., and
XX CC cardiovascular disorders such as cardiomyopathy, myocarditis, Duchenne's
XX CC muscular dystrophy, tachyarrhythmia etc. Antagonists and vectors
XX CC comprising antisense sequences are useful for treating and preventing
XX CC cancers such as adenocarcinoma, leukaemia, lymphoma and melanoma
XX SQ Sequence 147 AA;

Query Match 100.0%; Score 741; DB 3; Length 147;
Best Local Similarity 100.0%; Pred. No. 6.6e-68;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGGWHSQDQKSLHLEGDNPSPAPTSTCAPRMPKRIISIKOLASVKALRCSDEKAI 60
Db 1 MGGWHSQDQKSLHLEGDNPSPAPTSTCAPRMPKRIISIKOLASVKALRCSDEKAI 60
Qy 61 ATTALIFRNSSDSGKLEKAIADKLLQTFRNPFAEQETPKPKYREILSELDEHTENKLD 120
Db 61 ATTALIFRNSSDSGKLEKAIADKLLQTFRNPFAEQETPKPKYREILSELDEHTENKLD 120
Qy 121 EDFMILLSTIVMSDLLQNRNVKIMK 147
Db 121 EDFMILLSTIVMSDLLQNRNVKIMK 147

RESULT 2
AAW40026
ID AAW40026 standard; protein; 143 AA.
XX AC AAW40026;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 3171.
XX KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia.
XX OS Homo sapiens.
XX PN WO20015312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX DR N-PSDB; AA159182.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX PT as central nervous system injuries.
XX PS Example 4; SEQ ID NO 3171; 10078pp; English.
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and the
XX CC encoded polypeptides (AA138642-AA142213) with neurotropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders. Note: The sequence data for this patent did not form
XX CC part of the printed specification
XX SQ Sequence 143 AA;

Query Match 96.5%; Score 715; DB 4; Length 143;
Best Local Similarity 100.0%; Pred. No. 3e-65;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 MHSTQDKSLHLEGDNPSPAPTSTCAPRMPKRIISIKOLASVKALRCSDEKAIATTA 64
Db 1 MHSTQDKSLHLEGDNPSPAPTSTCAPRMPKRIISIKOLASVKALRCSDEKAIATTA 60
Qy 65 LIFRNSSDSGKLEKAIADKLLQTFRNPFAEQETPKPKYREILSELDEHTENKLD 124
Db 61 LIFRNSSDSGKLEKAIADKLLQTFRNPFAEQETPKPKYREILSELDEHTENKLD 120
Qy 125 ILLLSITVMSDLLQNRNVKIMK 147
Db 121 ILLLSITVMSDLLQNRNVKIMK 143

RESULT 3
AAU16494
ID AAU16494 standard; protein; 140 AA.
XX AC AAU16494;
XX DT 07-NOV-2001 (first entry)
XX DE Human novel secreted protein, Seq ID 1447.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
XX KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
XX KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;

KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX Homo sapiens.
XX WO200155322-A2.
PD 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001341.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205151P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0231415P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 03-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26481.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1447; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathologic condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 93.1%; Score 690; DB 4; Length 140;
Best Local Similarity 99.3%; Pred. No. 1.1e-62; Indels 0; Gaps 0;
Matches 138; Conservative 1; Mismatches 0;
QY 9 QDKSLHLEGDNPNSAAPTGTCAPRKPKKISISKQLASVKALKRKCSDLEKAIATTALIFR 68
Db :|||||
2 EDKSLHLEGDNPNSAAPTGTCAPRKPKKISISKQLASVKALKRKCSDLEKAIATTALIFR 61
QY 69 NSSDSGKLEKAIKADLLQTFNFAEGQETPKYREILSELDEHTENKLDPEDFMILL 128
Db :|||||
62 NSSDSGKLEKAIKADLLQTFNFAEGQETPKYREILSELDEHTENKLDPEDFMILL 121
QY 129 SITVMSDLLQNIIRNVKIMK 147
Db 122 SITVMSDLLQNIIRNVKIMK 140
RESULT 4
ABU55563
ID ABU55563 standard; protein; 140 AA.
AC ABU55563;
XX
XX 18-MAR-2003 (first entry)
XX
XX Human novel polypeptide #650.
XX
XX Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
XX

KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
XX Homo sapiens.
XX
XX US2002132753-A1.
XX
XX 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 20-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73822.
XX

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.

PS Claim 11; SEQ ID NO 1447; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU554914-ABU55499 and
 CC ABU55748 represent human novel polypeptides of the invention

XX Sequence 140 AA;

Query Match 93.1%; Score 690; DB 6; Length 140;

Best Local Similarity 99.3%; Pred. No. 1.1e-62;

Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 QDKSLHLEGGDPNPSAAPTSTCAPRMPKRISISKQLASVKALKKCSLDLEKATATTAIPR 68

Db 2 EDKSLHLEGGDPNPSAAPTSTCAPRMPKRISISKQLASVKALKKCSLDLEKATATTAIPR 61

QY 69 NSSDSGDKLEKAIADLLQTFQFNFAEGQETPKYREILSELDEHTENKLDPEDFMILL 128

Db 62 NSSDSGDKLEKAIADLLQTFQFNFAEGQETPKYREILSELDEHTENKLDPEDFMILL 121

QY 129 SITVMSDLLQNRVKNK 147

Db 122 SITVMSDLLQNRVKNK 140

RESULT 5

AAW41812

ID AAW41812 standard; protein; 137 AA.

AC AAW41812;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6743.

KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US034263.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;

DR WPI; 2001-442253/47.

XX N-FSDS; AAI60968.

PT Novel nucleic acids and polypeptides, useful for treating disorders such
 as central nervous system injuries.

XX Example 2; SEQ ID NO 6743; 10078pp; English.

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAW38642-AAW42213) with nontropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-prager Syndrome. Other uses include the
 CC utilisation of the activities such as; Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX Sequence 137 AA;

Query Match 92.6%; Score 686; DB 4; Length 137;

Best Local Similarity 100.0%; Pred. No. 2.7e-62;

Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 MHSTQDKSLHLEGGDPNPSAAPTSTCAPRMPKRISISKQLASVKALKKCSLDLEKAIATTA 64

Db 1 MHSTQDKSLHLEGGDPNPSAAPTSTCAPRMPKRISISKQLASVKALKKCSLDLEKAIATTA 60

QY 65 LIFRNSSDSGKLEKAIADLLQTFQFNFAEGQETPKYREILSELDEHTENKLDPEDFM 124

Db 61 LIFRNSSDSGKLEKAIADLLQTFQFNFAEGQETPKYREILSELDEHTENKLDPEDFM 120

QY 125 ILLLSITVMSDLLQNR 141

Db 121 ILLLSITVMSDLLQNR 137

RESULT 6

ADE61762

ID ADE61762 standard; protein; 93 AA.

AC ADE61762;

DT 29-JAN-2004 (first entry)

DE Rat Protein AAB20539, SEQ ID NO 7689.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

XX 27-FEB-2003.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GCHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 XX GENBANK; AAB20539.
 DR New composition comprising two or more isolated polypeptides, useful for
 DR preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 93 AA;

Query Match 18.7%; Score 138.5; DB 7; Length 93;
 Best Local Similarity 36.6%; Pred. No. 2.5e-06;
 Matches 30; Conservative 21; Mismatches 28; Indels 3; Gaps 1;
 Qy 54 SDLEKAIATTAIFRNSSDSG---KLEKAIADLLQTFNFAEQGQTKPKYREILSEL 110
 Db 2 SELETAMETLINVFHNSGEGDKYKLSKKEKLLQTELSSFLDQKADAVDKIMKEL 61
 Qy 111 DEHTENKLDPEDFMILLSTIV 132
 Db 62 DENGGEVDFOEYIVLVAALTU 83

RESULT 7
 ADE61764
 ID ADE61764 standard; protein; 93 AA.
 XX ADE61764;
 AC ADE61764;
 XX 29-JAN-2004 (first entry)
 DT Human Protein P23297, SEQ ID NO 7691.
 DE Human; pain; neuronal tissue; gene therapy;
 XX

KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS WO2003016475-A2.
 PN 27-FEB-2003.
 XX 14-AUG-2002; 2002WO-US025765.
 PF 14-AUG-2001; 2001US-0312147P.
 XX 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GCHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 XX GENBANK; P23297.
 DR New composition comprising two or more isolated polypeptides, useful for
 DR preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 93 AA;

Query Match 17.7%; Score 131.5; DB 7; Length 93;
 Best Local Similarity 32.9%; Pred. No. 1.3e-05;
 Matches 27; Conservative 23; Mismatches 29; Indels 3; Gaps 1;
 Qy 54 SDLEKAIATTAIFRNSSDSG---KLEKAIADLLQTFNFAEQGQTKPKYREILSEL 110
 Db 2 SELETAMETLINVFHNSGEGDKYKLSKKEKLLQTELSSFLDQKADAVDKIMKEL 61
 Qy 111 DEHTENKLDPEDFMILLSTIV 132
 Db 62 DENGGEVDFOEYIVLVAALTU 83

RESULT 8
 AAB45531

ID AAB45531 standard; protein; 94 AA.
 AC AAB45531;
 XX
 DT 22-FEB-2001 (first entry)
 XX
 DE Human S100A1 protein.
 XX
 KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
 KW calcium-binding protein; calcium homeostasis; cardiac muscle;
 KW pumping capacity; myocardial cell; systolic calcium ion release;
 KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
 KW valve defect.
 XX
 OS Homo sapiens.
 XX
 PN DE19915485-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-1999; 99DE-01015485.
 XX
 PR 07-APR-1999; 99DE-01015485.
 XX
 PA (KATU/) KATUS H A.
 PA (REMP/) REMPPIS A.
 XX
 PI Katus HA, Remppis A;
 XX
 PR WPI; 2000-673510/66.
 DR N-PSDB; AAC81801.
 DR
 XX
 PT Composition containing S100 protein, corresponding nucleic acid or
 PT vector, useful for treating cardiomyopathy and cardiac insufficiency.
 XX
 PS Claim 10; Page 8; 36pp; German.
 XX
 CC This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by
 CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (I) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease
 XX
 SQ Sequence 94 AA;
 Query Match 17.7%; Score 131.5; DB 3; Length 94;
 Best Local Similarity 32.9%; Pred. No. 1.3e-05;
 Matches 27; Conservative 23; Mismatches 29; Indels 3; Gaps 1;
 QY 54 SDLEKAIATTAIFRNSDSG---KLEKAKADLLQTOFRNFAEQGTQPKYREILSEL 110
 DB 3 SELETAMETLINVFHAKSGEGDKYLSKELKELLQTELSGFLDAQKDVAQDKWKEL 62
 QY 111 DEHTENKLDFFEDFMILLISITV 132
 DB 63 DENGSGEVDFOBYVVLVAALT 84
 RESULT 9

AAM40258
 ID AAM40258 standard; protein; 94 AA.
 XX
 AC AAM40258;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 3403.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00682191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI59414.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 5; SEQ ID NO 3403; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAI42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 94 AA;
 Query Match 17.7%; Score 131.5; DB 4; Length 94;
 Best Local Similarity 32.9%; Pred. No. 1.3e-05;
 Matches 27; Conservative 23; Mismatches 29; Indels 3; Gaps 1;
 QY 54 SDLEKAIATTAIFRNSDSG---KLEKAIKADLLQTOFRNFAEQGTQPKYREILSEL 110
 DB 3 SELETAMETLINVFHAKSGEGDKYLSKELKELLQTELSGFLDAQKDVAQDKWKEL 62

RESULT 14

ABG96408
ID ABG96408 standard; protein; 97 AA.

XX AC ABG96408;

XX DT 11-DEC-2002 (first entry)

XX DE Human ovarian cancer marker M68.

XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;
XX central nervous system disorder; bacterial meningitis; viral meningitis;
XX Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
XX brain herniation; inflammation; encephalitis; testicular disorder;
XX nontuberculous granulomatous orchitis; connective tissue disorder;
XX heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
XX histological type; carcinogenic; ovarian cancer marker.

XX OS Homo sapiens.

XX PN WO200271928-A2.

XX PD 19-SEP-2002.

XX PF 14-MAR-2002; 2002WO-US007826.

XX PR 14-MAR-2001; 2001US-0276025P.

XX PR 14-MAR-2001; 2001US-0276026P.

XX PR 10-AUG-2001; 2001US-0311732P.

XX PR 19-SEP-2001; 2001US-0323580P.

XX PR 26-SEP-2001; 2001US-0324967P.

XX PR 26-SEP-2001; 2001US-0325102P.

XX PR 26-SEP-2001; 2001US-0325149P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;

XX PI Meyers RE, Morrissey MP, Clandt PJ, Sen A, Vieby PO, Mills GS;

XX PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;

XX WPI; 2002-723277/78.

XX DR N-PSDB; ABS76507.

XX PT Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.

XX PS Disclosure; Page 392; 481pp; English.

XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX characterising cancer, in detecting the presence of cancer as early as
XX possible, and the recurrence of ovarian cancer. The method may also be of
XX particular use with patients having an enhanced risk of developing
XX ovarian cancer (e.g. patients having a familial history of ovarian
XX cancer). The cancer markers may be used in the management and treatment
XX of e.g. brain and central nervous system disorders (e.g. bacterial and
XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX testicular disorders (e.g. nontuberculous granulomatous orchitis),
XX connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX disease or atherosclerosis). The compositions and methods may also be
XX used in assessing the histological type of neoplasm associated with
XX ovarian cancer, monitoring the progression of ovarian cancer, determining
XX whether ovarian cancer has metastasized or is likely to metastasize,

CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention

XX SQ Sequence 97 AA;

Query Match 16.6%; Score 123; DB 5; Length 97;

Best Local Similarity 31.1%; Pred. No. 0.0001;

Matches 28; Conservative 22; Mismatches 36; Indels 4; Gaps 2;

QY 53 CSDLEKALATTALLFRNSSSDG--KLEKAIKDLLOTRNFAEQGQETPKKYRETLSE 109

DB 2 CSSLEQALAVLVITTFHKYSCQEGDKFKLSGEMKELLKELPSFGVKVDEGLKLMGS 61

QY 110 LDEHTENKLPEDFMILLISITVM-SDLLQ 138

DB 62 LDNSDQQVDFQYAVFLALITVMCNDFQ 91

RESULT 15

ABR92148

ID ABR92148 standard; protein; 97 AA.

XX AC ABR92148;

XX DT 10-SEP-2003 (first entry)

XX DE Human cervical cancer cell marker protein SEQ ID NO:206.

XX Human; cervical cancer; cervical cancer marker; cancer therapy;

XX detection; gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO2002101075-A2.

XX PD 19-DEC-2002.

XX PF 12-JUN-2002; 2002WO-US018638.

XX PR 13-JUN-2001; 2001US-0298155P.

XX PR 13-JUN-2001; 2001US-0298159P.

XX PR 14-NOV-2001; 2001US-0335936P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;

XX PI Gannavarapu M, Glatt K, Hoersch S;

XX WPI; 2003-156967/15.

XX DR N-PSDB; ACF12931.

XX New isolated nucleic acid molecule useful for detecting, characterizing,
XX preventing and treating human cervical cancers, in various prognostic and
XX diagnostic assays, in pharmacogenomics and in monitoring clinical trials.

XX Claim 4; Page 361-362; 386pp; English.

XX ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
XX given in ABR92047 to ABR92164. A higher level of expression of (I) than
XX normal indicates the presence of cervical cancer. Also described: (1) a
XX vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
XX assessing (M) whether a patient is afflicted with cervical cancer,
XX comprising comparing the level of expression of a marker in a patient's
XX sample, and the normal level of expression of the marker in a control non
XX cervical cancer sample, where a significant increase in the level of
XX expression of the marker in the patient's sample relative to that in the
XX control sample is an indication that the patient is afflicted with
XX cervical cancer. (I) has cytostatic activity, and can be used in gene
XX therapy and in vaccines. (I) is useful in detecting, characterising,
XX preventing and treating human cervical cancers. (I) may also be used in

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:26:27 ; Search time 40 Seconds
(without alignments)
1159.529 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 147
Sequence: 1 MGGCMHSTQDKSLHLEDPN.....LSITWMSDLLQNIENVKIMK 147

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	12.2	147	11 Q8C9X1	Q8C9X1 mus musculus
2	9	6.1	360	10 Q940C2	Q940C2 arabidopsis
3	8	5.4	314	10 P93444	P93444 pyroplasma
4	8	5.4	424	11 Q99N63	Q99N63 mesocricetu
5	8	5.4	716	2 Q49526	Q49526 mycoplasma
6	8	5.4	730	10 Q9SKK7	Q9SKK7 arabidopsis
7	8	5.4	1051	2 Q49524	Q49524 mycoplasma
8	8	5.4	1344	2 Q49545	Q49545 mycoplasma
9	8	5.4	1365	2 Q49525	Q49525 mycoplasma
10	8	5.4	1860	5 Q9VYP5	Q9VYP5 drosophila
11	7	4.8	111	8 Q94RG1	Q94RG1 heterodoncu
12	7	4.8	111	16 Q8YKY3	Q8YKY3 anabaena sp
13	7	4.8	114	16 Q7URH0	Q7URH0 rhodospirell
14	7	4.8	116	8 Q94ZD6	Q94ZD6 eptatretus
15	7	4.8	130	16 Q7V175	Q7V175 prochloroco
16	7	4.8	131	11 Q8BYN5	Q8BYN5 mus musculu

17	7	4.8	145	4 Q9BXM9	Q9BXM9 homo sapien
18	7	4.8	156	16 Q8CS56	Q8CS56 staphylococ
19	7	4.8	157	16 Q7UGD3	Q7UGD3 rhodospirell
20	7	4.8	156	10 Q9ZRF2	Q9ZRF2 nicotiana t
21	7	4.8	172	13 Q13095	Q13095 salmo salar
22	7	4.8	177	10 Q9FXT1	Q9FXT1 nicotiana t
23	7	4.8	179	5 Q9VD02	Q9VD02 drosophila
24	7	4.8	185	16 Q67526	Q67526 aquifex aeo
25	7	4.8	214	10 Q41899	Q41899 zea mays (m
26	7	4.8	220	17 Q974S7	Q974S7 sulfolobus
27	7	4.8	234	1 Q72AH4	Q72AH4 haloarcula
28	7	4.8	237	16 Q8UI09	Q8UI09 agrobacteri
29	7	4.8	247	10 Q40900	Q40900 petunia int
30	7	4.8	247	10 Q08711	Q08711 petunia hyb
31	7	4.8	259	10 Q24009	Q24009 zea mays (m
32	7	4.8	259	16 Q9K5U6	Q9K5U6 bacillus ha
33	7	4.8	260	2 Q8GDY3	Q8GDY3 helicobacill
34	7	4.8	264	10 Q9ZS30	Q9ZS30 getbera hyb
35	7	4.8	265	16 Q8DV40	Q8DV40 streptococc
36	7	4.8	269	10 Q8H281	Q8H281 streptococc
37	7	4.8	273	10 Q84V80	Q84V80 zea mays (m
38	7	4.8	277	2 Q8VUN3	Q8VUN3 streptococc
39	7	4.8	291	17 Q8TFS8	Q8TFS8 methanosarc
40	7	4.8	295	16 Q9A2A7	Q9A2A7 caulobacter
41	7	4.8	295	16 Q894V2	Q894V2 clostridium
42	7	4.8	304	10 Q84V75	Q84V75 zea mays (m
43	7	4.8	309	16 Q89LT0	Q89LT0 bradyrhizob
44	7	4.8	324	16 Q88Z06	Q88Z06 lactobacill
45	7	4.8	341	8 Q8WE39	Q8WE39 draco taeni

ALIGNMENTS

RESULT 1					
Q8C9X1	PRELIMINARY;	PRT;	147 AA.		
ID	Q8C9X1				
AC	Q8C9X1; 2003 (TRENBLrel. 23, Created)				
DT	01-MAR-2003 (TRENBLrel. 23, Last sequence update)				
DT	01-MAR-2003 (TRENBLrel. 25, Last annotation update)				
DE	Hypothetical EF-hand containing protein.				
OS	Mus musculus (Mouse)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Thymus;				
RX	MEDLINE=22354683; PubMed=12466851;				
RA	The FANTOM Consortium,				
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;				
RT	"Analysis of the mouse transcriptome based on functional annotation of				
RT	60,770 full-length cDNAs."				
RL	Nature 420:563-573(2002).				
DR	EMBL; AK040281; BAC30559.1; "				
DR	GO; GO:0005509; F;calcium ion binding; IEA.				
DR	InterPro; IPR002048; EF-hand.				
KW	Hypothetical protein.				
SQ	SEQUENCE 147 AA; 16415 MW; FICE14DBFEB26D59 CRC64;				
Query Match	12.2%; Score 18; DB 11; Length 147;				
Best Local Similarity	100.0%; Pred. No. 2.9e-10;				
Matches 18; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	106 ILSELDHTEKNKLDPEDF 123				
Db	106 ILSELDHTEKNKLDPEDF 123				
RESULT 2					
Q940C2	PRELIMINARY;	PRT;	360 AA.		
ID	Q940C2				

AC Q940C2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE A759250/f6N7_3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RN Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY056121; AAL07007.1; -;
DR EMBL; AY078034; AAL77735.1; -;
SQ SEQUENCE 360 AA; 41042 MW; E5991D998FF49F01 CRC64;

Query Match 6.1%; Score 9; DB 10; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LEXA1ATTA 64
Db 143 LEXA1ATTA 151

RESULT 3
ID P93444 PRELIMINARY; PRT; 314 AA.
AC P93444;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P0kweed antiviral protein precursor (BC 3.2.2.22) (rRNA
DE N-glycosidase).
GN PAP-S.
OS Phytolacca americana (Common pokeberry) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Phytolaccaceae; Phytolacca.
OX NCBI_TaxID=3527;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Seed;
RX MEDLINE=97263479; PubMed=9109394;
RA Poyet J.L., Hoeweler A.;
RT "cDNA cloning of the gene encoding the Antiviral Protein from the
RT seeds of Phytolacca americana and its expression in E.coli.";
RL FEBS Lett. 406:97-100 (1997).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; X98079; CRA66702.1; -;

DR HSP; Q03464; IAPA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIF.
DR Pfam; PF00161; RIF; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 24
FT CHAIN 25 314 POTENTIAL.
SQ SEQUENCE 314 AA; 35323 MW; A89B3CE57789FF9E CRC64;

Query Match 5.4%; Score 8; DB 10; Length 314;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AAPTSTCA 30
Db 17 AAPTSTCA 24

RESULT 4
ID Q99N63 PRELIMINARY; PRT; 424 AA.
AC Q99N63;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeobox B3.
GN HOX B3.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RN SEQUENCE FROM N.A.
RA Nakamura N.;
RT "hamster Hox B3 cds.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AB056577; BAB33375.1; -;
DR HSP; P02833; 18AN.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00027; HOMEBOX; 1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 424 AA; 43631 MW; DC7E64040B8DA5A4 CRC64;

Query Match 5.4%; Score 8; DB 11; Length 424;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 PSAAPTST 28
Db 92 PSAAPTST 99

RESULT 5
Q49526

```
ID Q49526 PRELIMINARY; PRT; 716 AA.
AC Q49526;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Lmp1.
GN LMP1.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wh74;
RX MEDLINE=95369882; PubMed=7543881;
RA Jensen L.T., Ladefoged S., Birkelund S., Christiansen G.;
RT "Selection of Mycoplasma hominis PG21 deletion mutants by cultivation
in the presence of monoclonal antibody 552."
RL Infect. Immun. 63:3336-3347(1995).
DR EMBL; U21963; AAA81014.1; -.
DR InterPro; IPR006864; LMP.
DR Pfam; PF04778; LMP; 4.
SQ SEQUENCE 716 AA; 79864 MW; 28D08C3E0C91CB47 CRC64;

Query Match 5.4%; Score 8; DB 2; Length 716;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 AIATTALI 66
DB 20 AIATTALI 27

RESULT 6
Q3SKK7 PRELIMINARY; PRT; 730 AA.
ID Q3SKK7;
AC Q3SKK7;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN AT2G25420
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006300; RAD20702.2; -.
DR FIR; B84648; B84648.
DR InterPro; IPR006595; CTLH.
DR InterPro; IPR006594; LISH.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 2.
DR SMART; SM00668; CTLH; 2.
DR SMART; SM00667; LISH; 2.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS00897; CTLH; 1.
DR PROSITE; PS00896; LISH; 2.
DR PROSITE; PS00882; WD REPEATS 2; 1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Hypothetical protein_Repeat; WD repeat.
SQ SEQUENCE 730 AA; 82003 MW; 18205CE1D3769F88 CRC64;

Query Match 5.4%; Score 8; DB 10; Length 730;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 70 SSDSGKL 77
DB 541 SSDSGKL 548

RESULT 7
Q49524 PRELIMINARY; PRT; 1051 AA.
ID Q49524;
AC Q49524;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Lmp1.
GN LMP1.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wh74;
RX MEDLINE=95369882; PubMed=7543881;
RA Jensen L.T., Ladefoged S., Birkelund S., Christiansen G.;
RT "Selection of Mycoplasma hominis PG21 deletion mutants by cultivation
in the presence of monoclonal antibody 552."
RL Infect. Immun. 63:3336-3347(1995).
DR EMBL; U21961; AAA81012.1; -.
DR FIR; T18351; T18351.
DR InterPro; IPR006864; LMP.
DR Pfam; PF04778; LMP; 2.
SQ SEQUENCE 1051 AA; 120260 MW; 82B311BB966BFF7E CRC64;
```

```
Query Match 5.4%; Score 8; DB 2; Length 1051;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 59 AIATTALI 66
DB 20 AIATTALI 27

RESULT 8
Q49545 PRELIMINARY; PRT; 1344 AA.
ID Q49545;
AC Q49545;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Lmp1 protein.
GN LMP1.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG21;
RX MEDLINE=95104998; PubMed=7806360;
RA Ladefoged S.A., Birkelund S., Rauge S., Brock B., Jensen L.T.,
RA Christiansen G.;
RT "A 135-kilodalton surface antigen of Mycoplasma hominis PG21 contains
multiple directly repeated sequences."
RL Infect. Immun. 63:212-223(1995).
DR EMBL; X81475; CAA57228.1; -.
DR InterPro; IPR006864; LMP.
DR Pfam; PF04778; LMP; 8.
SQ SEQUENCE 1344 AA; 149405 MW; 6E0105F7365AF8CF CRC64;
```

```
Query Match 5.4%; Score 8; DB 2; Length 1344;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 59 AIATTALI 66
DB 20 AIATTALI 27

RESULT 8
Q49545 PRELIMINARY; PRT; 1344 AA.
ID Q49545;
AC Q49545;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Lmp1 protein.
GN LMP1.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG21;
RX MEDLINE=95104998; PubMed=7806360;
RA Ladefoged S.A., Birkelund S., Rauge S., Brock B., Jensen L.T.,
RA Christiansen G.;
RT "A 135-kilodalton surface antigen of Mycoplasma hominis PG21 contains
multiple directly repeated sequences."
RL Infect. Immun. 63:212-223(1995).
DR EMBL; X81475; CAA57228.1; -.
DR InterPro; IPR006864; LMP.
DR Pfam; PF04778; LMP; 8.
SQ SEQUENCE 1344 AA; 149405 MW; 6E0105F7365AF8CF CRC64;
```

```
Query Match 5.4%; Score 8; DB 2; Length 1344;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 59 AIATTALI 66
DB 20 AIATTALI 27
```

```

Db      20 A1ATTALI 27

RESULT 9
Q49525 PRELIMINARY; PRT; 1365 AA.
AC Q49525;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Lmp1.
GN Lmp1.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mh56;
RX MEDLINE=95369882; PubMed=7543881;
RA Jensen L.T., Ladefoged S., Birkelund S., Christiansen G.;
RT "Selection of Mycoplasma hominis PG21 deletion mutants by cultivation
in the presence of monoclonal antibody 552.";
RL Infect. Immun. 63:3336-3347(1995).
DR EMBL; U21962; AAA81013.1; -.
DR PIR; T30822; T30822.
DR InterPro; IPR006864; LMP.
DR Pfam; PF04778; LMP; 4.
SQ SEQUENCE 1365 AA; 154983 MW; 847A04992410867F CRC64;

Query Match 5.4%; Score 8; DB 2; Length 1365;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 A1ATTALI 66
Db 20 A1ATTALI 27

RESULT 10
Q9VYPS PRELIMINARY; PRT; 1860 AA.
AC Q9VYPS;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE CG2750 protein.
GN CG2750.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Ephydroidea; Drosophilidae; Diptera; Brachycera; Muscomorpha;
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC Evans S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Celniker S.E., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak P., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE003488; AAF48147.2; -.
DR FlyBase; FBN0030376; CG2750.
DR InterPro; IPR004325; Nucleoporin.FG.
DR Pfam; PF03093; Nucleoporin.FG; 10.
SQ SEQUENCE 1860 AA; 206779 MW; 04935971B7141272 CRC64;

Query Match 5.4%; Score 8; DB 5; Length 1860;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 ISISKOLA 45
Db 1692 ISISKOLA 1699

RESULT 11

```

```
Q94RGI
ID Q94RGI1 PRELIMINARY; PRT; 111 AA.
AC
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE NADH subunit 3 (Fragment).
GN NADH3.
OS Heterodontus francisci (Horn shark).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OX NCBI_TaxID=7792;
RN [1]_TaxID=7792;
RP SEQUENCE FROM N.A.
RA Arnason U., Gullberg A., Janke A.;
RT "Molecular phylogenetics of gnathostomous (jawed) fishes: Old bones,
RT new cartilage.";
RL Zool. Scr. 30:249-255 (2001).
DR EMBL; AJ310141; CAC94219.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR000440; Oxidored_q4.
DR Pfam; PF00507; Oxidored_q4; 1.
KW Mitochondrion.
FT NON TER. 111
SQ SEQUENCE 111 AA; 12355 MW; D670DA59C28772EC CRC64;

Query Match 4.8%; Score 7; DB 8; Length 111;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 IATTALI 66
DB 8 IATTALI 14

RESULT 12
Q8YKX3 PRELIMINARY; PRT; 111 AA.
ID Q8YKX3
AC
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein All7155.
GN All7155.
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120alpha.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213 (2001).
DR EMBL; AF003600; BAB78239.1; -.
DR PIR; AC2497; AC2497.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR InterPro; IPR007712; Plasmid stabil.
DR Pfam; PF05016; Plasmid stabil; 1.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 111 AA; 12855 MW; C474F9277DE8A3AL CRC64;

Query Match 4.8%; Score 7; DB 16; Length 111;
Best Local Similarity 100.0%; Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 IATTALI 66
DB 8 IATTALI 14

RESULT 13
Q7URHO PRELIMINARY; PRT; 114 AA.
ID Q7URHO
AC Q7URHO;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Probable molybdopter-in-synthase sulfurylase.
GN MOEB OR RB5670.
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."; Acad. Sci. U.S.A. 100:8298-8303 (2003).
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294142; CAD74368.1; -.
KW Complete proteome.
SQ SEQUENCE 114 AA; 12973 MW; 4627CA74011C344D CRC64;

Query Match 4.8%; Score 7; DB 16; Length 114;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LSELDEH 113
DB 57 LSELDEH 63

RESULT 14
Q94ZD6 PRELIMINARY; PRT; 116 AA.
ID Q94ZD6
AC Q94ZD6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 3.
GN NADH3.
OS Eptatretus burgeri (Inshore hagfish).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptaretinae; Eptatretus.
OX NCBI_TaxID=7764;
RN [1]_
RP SEQUENCE FROM N.A.
RA Delarbre C., Gallut C., Barriel V., Janvier P., Gachelin G.;
RT "Phylogenetic resolution of cyclostomes using the mitochondrial DNA
RT sequence of the hagfish Eptatretus burgeri.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278504; CAC42109.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR000440; Oxidored_q4.
DR Pfam; PF00507; Oxidored_q4; 1.
KW Mitochondrion.
SQ SEQUENCE 116 AA; 13277 MW; BFF4F1AD9297233B CRC64;

Query Match 4.8%; Score 7; DB 8; Length 116;
Best Local Similarity 100.0%; Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ILSELDE 112
DB 103 ILSELDE 109
```


Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 ILLLSIT 131
Db 5 ILLLSIT 11

RESULT 15

Q7V175
ID Q7V175 PRELIMINARY; PRT; 130 AA.
AC Q7V175;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical precursor.
GN PM1008.
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.,
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation."
RL Nature 424:1042-1047(2003).
DR EMBL; BX572092; CAB19467.1; --
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 23
SQ SEQUENCE 130 AA; 14889 MW; 0D8681ADCAF81E77 CRC64;

Query Match 4.8%; Score 7; DB 16; Length 130;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 MILLLSI 130
Db 1 MILLLSI 7

Search completed: June 3, 2004, 13:30:21
Job time : 41 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:23:07 ; Search time 18 Seconds
(without alignments)
425.240 Million cell updates/sec

Title: US-09-744-197-1

Perfect score: 147

Sequence: 1 MGCCHSTQKSLHLEGP.....LSITWSDLLQNIIRNVKIMK 147

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	8	5.4	433	1 HXB3_MOUSE	P09026 mus musculus
2	7	4.8	161	1 LEPA_PEEFL	P26843 pseudomonas
3	7	4.8	182	1 ARF3_YEAST	P40994 saccharomyc
4	7	4.8	322	1 RL40_LUPLU	P50345 lupinus lut
5	7	4.8	346	1 PER9_ARATH	Q96512 arabidopsis
6	7	4.8	370	1 LEU3_RHIME	Q92488 rhizobium m
7	7	4.8	370	1 PRIL_AECFU	Q29911 archaeglob
8	7	4.8	429	1 YHG3_YEAST	P38756 saccharomyc
9	7	4.8	431	1 HXB3_HUMAN	P14651 homo sapien
10	7	4.8	439	1 PROA_SYNEL	Q8dku1 synechococc
11	7	4.8	459	1 GATE_TRETH	Q91cx2 thermus the
12	7	4.8	497	1 C71G_ARATH	Q9fh66 arabidopsis
13	7	4.8	504	1 SH2P_MOUSE	Q02152 mus musculus
14	7	4.8	519	1 SYH_FUGRU	P70076 fugu rubrip
15	7	4.8	573	1 DPA1_SYNY3	Q55393 synechocyst
16	7	4.8	614	1 LEPA_CXBU	Q83bk3 coxiella bu
17	7	4.8	618	1 CHEA_LISIN	Q92dw2 listeria in
18	7	4.8	618	1 CHEA_LISMO	Q48768 listeria mo
19	7	4.8	813	1 ARM_MUSDO	Q02453 musca domes
20	7	4.8	980	1 UBEA_CABEL	Q03349 xenorhabdi
21	7	4.8	1807	1 UTA2_XENLA	P18709 xenopus lae
22	7	4.8	3023	1 POLG_TVMV	P09814 t genome po
23	6	4.1	44	1 PSBN_CHLVU	P56326 chlorella v
24	6	4.1	64	1 RL37_LYCES	P49212 lycopersico
25	6	4.1	65	1 ICE2_ASCSU	P07852 ascaris suu
26	6	4.1	65	1 PAPE_PHYAM	P81418 phytolacca
27	6	4.1	69	1 COXR_PAPHA	Q8sq78 papio hamad
28	6	4.1	72	1 RL15_BACLI	P35138 bacillus li
29	6	4.1	78	1 Y270_METJA	Q57718 methanococc
30	6	4.1	92	1 FNR_CAPOC	Q46158 capnocytoph
31	6	4.1	92	1 FNR_CARSP	Q46159 capnocytoph
32	6	4.1	92	1 FNR_HAEP	Q47948 haemophilus
33	6	4.1	93	1 S10A_MOUSE	P56565 mus musculus

34	6	4.1	93	1 S10A_RAT	P35467 rattus norv
35	6	4.1	97	1 HFO_NEIMA	Q9jgw6 neisseria m
36	6	4.1	100	1 YNO5_YEAST	P53906 saccharomyc
37	6	4.1	102	1 YE07_METH	Q27458 methanobact
38	6	4.1	110	1 LZ8_GANLU	P14945 ganoderma l
39	6	4.1	112	1 UCN2_HUMAN	Q96tp3 homo sapien
40	6	4.1	120	1 RL18_TREPA	O83335 treponema p
41	6	4.1	128	1 COTV_BACSU	Q08309 bacillus su
42	6	4.1	129	1 MINK_HUMAN	P15382 homo sapien
43	6	4.1	134	1 ATPE_DSVM	Q985b5 desulfovibr
44	6	4.1	136	1 L5E2_HUMAN	O43716 homo sapien
45	6	4.1	138	1 Y560_METJA	Q57980 methanococc

ALIGNMENTS

RESULT 1

ID	HXB3_MOUSE	STANDARD;	PRT;	433 AA.
AC	P09026; P10285; Q61680;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Homeobox protein Hox-B3 (Hox-2.7) (MH-23).			
GN	HOXB3 OR HOXB-3 OR HOX-2.7.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92258392; PubMed=1582411;			
RA	Sham M.H., Hunt P., Nonchev S., Papalopulu N., Graham A.,			
RA	Boncinelli E., Krumlauf R.;			
RT	"Analysis of the murine Hox-2.7 gene: conserved alternative			
RT	transcripts with differential distributions in the nervous system and			
RT	the potential for shared regulatory regions.";			
RL	EMBO J. 11:1825-1836(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95196953; PubMed=7890121;			
RA	Brown W.M., Taylor G.R.;			
RA	"The 5'-sequence of the murine Hox-b3 (Hox-2.7) gene and its intron			
RT	contain multiple transcription-regulatory elements.";			
RL	Int. J. Biochem. 26:1403-1409(1994).			
RN	[3]			
RP	SEQUENCE OF 152-361 FROM N.A.			
RX	MEDLINE=88054465; PubMed=2890503;			
RA	Lonai P., Arman E., Czosnek H., Ruddle F.H., Blatt C.;			
RT	"New murine homeoboxes: structure, chromosomal assignment, and			
RT	differential expression in adult erythropoiesis.";			
RL	DNA 6:409-418(1987).			
RN	[4]			
RP	SEQUENCE OF 181-265 FROM N.A.			
RX	MEDLINE=89091992; PubMed=2463210;			
RA	Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,			
RA	Krumlauf R.;			
RT	"Characterization of a murine homeo box gene, Hox-2.6, related to the			
RT	Drosophila Deformed gene.";			
RL	Genes Dev. 2:1424-1438(1988).			
CC	-!- FUNCTION: Sequence-specific transcription factor which is part of			
CC	a developmental regulatory system that provides cells with			
CC	specific positional identities on the anterior-posterior axis.			
CC	-!- SUBCELLULAR LOCATION: Nuclear.			
CC	-!- SIMILARITY: Belongs to the Antp homeobox family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			

```

CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; X66177; CAA46951.1; -
CC ENBL; U02276; AAB60496.1; -
CC ENBL; M8168; AAA37840.1; -
CC PIR; S20963; S20963.
CC HSP; P02833; ISAN.
CC TRANSFAC; T01724; -
CC MGD; MGI:96184; Hoxb3.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC PRINTS; PR00046; Homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SMC0389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC SITE 129 134
CC DOMAIN 154 181
CC DNAS_BIND 191 250
CC FT CONFLICT 113 113
CC FT CONFLICT 119 119
CC FT CONFLICT 152 168
CC FT CONFLICT 182 182
CC FT CONFLICT 216 217
CC FT CONFLICT 330 330
CC FT CONFLICT 342 361
CC FT CONFLICT 433 433
CC SEQUENCE 433 AA; 44353 MW; 9AD3C922663612A6 CRC64;

Query Match 5.4%; Score 8; DB 1; Length 433;
Best Local Similarity 100.0%; Pred.No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 PSAAPTST 28
DB 92 PSAAPTST 99

RESULT 2
LEPA_PSEFL STANDARD; PRT; 161 AA.
AC P26843;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding protein lepa (Fragment).
GN LEPA.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 10586;
RA Black M.T., Munn J.G.R., Allsop A.E.;
RT "On the catalytic mechanism of prokaryotic leader peptidase 1.";
RL Biochem. J. 282:539-544(1992).
CC -!- SUBCELLULAR LOCATION: PRESENT IN THE CYTOPLASM (BY SIMILARITY).
CC ALSO FOUND IN THE PERIPLASM (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC Lepa subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; L28996; AAA61614.1; -
CC ENBL; X94335; CAA64016.1; -
CC ENBL; Z75002; CAA99291.1; -
CC PIR; A53917; A53917.
CC HSP; P32889; IRRP.
CC GerMOnline; 143682; -
CC SGD; S0005620; ARF3.

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; X56466; CAA39838.1; -
CC PIR; S36880; S36880.
CC HAMAP; MF_00071; -; 1.
CC InterPro; IPR000795; EF_GTPbind.
CC InterPro; IPR000640; EFG_C.
CC InterPro; IPR009022; EFG_III_V.
CC Pfam; PF00679; EFG_C; 1.
CC PROSITE; PS00301; EFATOR_GTP; PARTIAL.
CC Membrane; GTP-binding.
CC NON_TER 1
CC SEQUENCE 161 AA; 18210 MW; 5FB6924B6C47BB50 CRC64;

Query Match 4.8%; Score 7; DB 1; Length 161;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 SVKALRK 52
DB 106 SVKALRK 112

RESULT 3
ARF3_YEAST STANDARD; PRT; 182 AA.
AC P40994;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADP-ribosylation factor 3.
GN ARF3 OR YOR094W OR YOR3172W
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94342251; PubMed=8063710;
RA Lee F.-J.S., Stevens L.A., Kao Y.L., Moss J., Vaughan M.;
RT "Characterization of a glucose-repressible ADP-ribosylation factor 3
(RARF3) from Saccharomyces cerevisiae."
RL J. Biol. Chem. 269:20931-20937(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97344369; PubMed=9200815;
RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.;
RA Schwaiger C., Paces V., Sander C., Ansorge W.;
RT "DNA sequencing and analysis of 130 Kb from Yeast chromosome XV."
RL Yeast 13:655-672(1997).
CC -!- FUNCTION: GTP-binding protein that functions as an allosteric
CC activator of the cholera toxin catalytic subunit, an ADP-
CC ribosyltransferase. Involved in protein trafficking; may modulate
CC vesicle budding and uncoating within the Golgi apparatus.
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Arf family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; L28996; AAA61614.1; -
CC ENBL; X94335; CAA64016.1; -
CC ENBL; Z75002; CAA99291.1; -
CC PIR; A53917; A53917.
CC HSP; P32889; IRRP.
CC GerMOnline; 143682; -
CC SGD; S0005620; ARF3.

```

DR InterPro; IPR006688; ARF.
DR InterPro; IPR006689; ARF/SAR.
DR InterPro; IPR001806; Ras transfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00025; arf; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR PRINTS; PR00328; SARIGTFBP.
DR SMART; SM00177; ARF; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS01019; ARF; 1.
KW GTP-binding; Multigene family; Myristate; Protein transport;
KW Golgi stack; Lipoprotein.
FT INTR_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 N-myristoyl glycine (Potential).
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
SQ SEQUENCE 182 AA; 20565 MW; 6B8370128369A270 CRC64;
Query Match 4.8%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 ATTALIF 67
DB 83 ATTALIF 89
RESULT 4
ID RLAO LUPLU STANDARD; PRT; 322 AA.
AC P50345;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60S acidic ribosomal protein P0.
OS Lupinus luteus (Yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
OX NCBI_TaxID=3873;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ventus;
RX MEDLINE=98163879; PubMed=9503149;
RA Mikolajczyk K., Barciszewski J.;
RT "Expression of the cDNA and purification of P0 ribosomal protein from
Lupinus luteus.";
RL Biochem. Mol. Biol. Int. 44:69-77(1998).
CC -!- FUNCTION: Ribosomal protein P0 is the functional equivalent of
E. coli protein L10.
CC -!- SUBUNIT: P0 forms a pentameric complex by interaction with dimers
of P1 and P2 (By similarity).
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X93587; CAA63786.1; -
DR InterPro; IPR001813; Ribosomal 60S.
DR InterPro; IPR001790; Ribosomal L10.
DR Pfam; PF00428; 60S ribosomal L10.
DR Pfam; PF00456; Ribosomal L10; 1.
KW Ribosomal protein; Phosphorylation.
SQ SEQUENCE 322 AA; 34453 MW; DA8566209B5FC0BC CRC64;
Query Match 4.8%; Score 7; DB 1; Length 322;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 PNPSAAP 25
DB 235 PNPSAAP 241
RESULT 5
ID PER9 ARATH STANDARD; PRT; 346 AA.
AC Q96512; Q9LPD4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Peroxidase 9 precursor (EC 1.11.1.7) (Atperox P9) (ATPI8a).
GN PER9 OR P9 OR At1G44970 OR F27F5.6 OR T22C22.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Welinder K.G., Jespersen H.M., Kjaersgaard I.V.H., Justesen A.F.,
Oestergaard L., Abelskov A.K., Hansen L.N., Rasmussen S.K.;
RT "From expressed sequence tags to structure, function, evolution and
expression of 29 ER-targeted Arabidopsis peroxidases.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Pederspiel N.A., Kaul S.,
White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Dewar K.,
Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
Gunter J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
Milticher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Brever V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
Feldmann K.A.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RC STRAIN=cv. Columbia;
RX MEDLINE=98409321; PubMed=9738941;
RA Oestergaard L., Pedersen A.G., Jespersen H.M., Brunak S.,
Welinder K.G.;
RT "Computational analyses and annotations of the Arabidopsis peroxidase
gene family.";
RL FEBS Lett. 433:98-102(1998).
RN [5]
RP GENE FAMILY ORGANIZATION, AND NOMENCLATURE.
RC STRAIN=cv. Columbia;
RX MEDLINE=22030461; PubMed=12034502;
RA Tognolli M., Penel C., Greppin H., Simon P.;
RT "Analysis and expression of the class III peroxidase large gene family

RT in Arabidopsis thaliana. ";
RL Gene 288:129-138(2002).
CC -!- FUNCTION: Removal of H(2)O(2), oxidation of toxic reductants,
CC biosynthesis and degradation of lignin, suberization, auxin
CC catabolism, response to environmental stresses such as wounding,
CC pathogen attack and oxidative stress. These functions might be
CC dependent on each isozyme/isoform in each plant tissue.
CC -!- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.
CC -!- COFACTOR: Binds heme B (iron-protoporphyrin IX) group and 2
CC calcium ions per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- MISCELLANEOUS: There are 73 peroxidase genes in A.thaliana.
CC -!- SIMILARITY: Belongs to the peroxidase family. Classical plant
CC (class III) peroxidase subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X59804; CAA67336.1; -;
CC EMBL; AC007915; AAF69153.1; -;
CC EMBL; AC020576; AAF78280.1; -;
CC EMBL; AY086626; AAM63684.1; -;
CC HSP; P22195; LSCH.
CC InterPro; IPR002016; Peroxidase.
CC Pfam; PF00141; peroxidase; 1.
CC PRINTS; PR00458; PEROXIDASE.
CC PROSITE; PS00435; PEROXIDASE_1; 1.
CC PROSITE; PS00436; PEROXIDASE_2; 1.
CC PROSITE; PS00873; PEROXIDASE_4; 1.
CC Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;
CC Multigene family; Calcium; Signal.
CC SIGNAL
FT CHAIN 1 23 POTENTIAL.
FT METAL 24 346 PEROXIDASE 9.
FT METAL 86 86 CALCIUM 1 (BY SIMILARITY).
FT METAL 89 89 CALCIUM 1 (VIA CARBONYL OXYGEN)
FT METAL 91 91 CALCIUM 1 (VIA CARBONYL OXYGEN)
FT METAL 91 91 CALCIUM 1 (BY SIMILARITY).
FT METAL 93 93 CALCIUM 1 (BY SIMILARITY).
FT METAL 95 95 CALCIUM 1 (BY SIMILARITY).
FT METAL 213 213 CALCIUM 2 (BY SIMILARITY).
FT METAL 264 264 CALCIUM 2 (BY SIMILARITY).
FT METAL 267 267 CALCIUM 2 (BY SIMILARITY).
FT METAL 272 272 CALCIUM 2 (BY SIMILARITY).
FT ACT_SITE 81 81 BY SIMILARITY.
FT ACT_SITE 85 85 DISTAL HISTIDINE.
FT ACT_SITE 113 113 HYDROGEN-BOUND (BY SIMILARITY).
FT ACT_SITE 182 182 SUBSTRATE BINDING (BY SIMILARITY).
FT METAL 212 212 IRON (HEME AXIAL LIGAND).
FT ACT_SITE 289 289 BY SIMILARITY.
FT ACT_SITE 54 134 HYDROGEN-BOUND (BY SIMILARITY).
FT DISULFID 87 92 BY SIMILARITY.
FT DISULFID 140 342 BY SIMILARITY.
FT DISULFID 219 251 BY SIMILARITY.
FT CARBOHYD 185 185 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 269 269 A -> S (IN REF. 3).
FT CONFLICT 290 290 E -> Q (IN REF. 3).
FT CONFLICT 302 302 A -> S (IN REF. 3).
SQ SEQUENCE 346 AA; 37741 MW; 47EFD6297469A065 CRC64;

Query Match 4.8%; Score 7; DB 1; Length 346;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 LEKAIK 83
| | | | |
Db 65 LEKAIK 71

RESULT 6
LEUS_RHIME STANDARD; PRT; 370 AA.
ID LEU3_RHIME
AC Q92KY8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE (IMDH) (3-IPM-DH).
GN LEUB OR R0310 OR SMC04405
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN (1)
RP SEQUENCE FROM N.A.
RC SFPAIN=1021.
RX MEDLINE=2136507; PubMed=11481430;
RA Capela D., Harloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kias E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ransperger U.,
RA Renard C., Thebault P., Vandebol M., Weidner S., Galibert F.;
RA "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-
CC methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-
CC oxopentanoate. The product decarboxylates to 4-methyl-2-
CC oxopentanoate.
CC -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
CC NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.
CC -!- PATHWAY: Leucine biosynthesis; third step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate
CC dehydrogenases family. LeuB subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL591793; CAC47889.1; -;
CC HAMAP; MF_01033; -; 1.
CC InterPro; IPR001804; Isodh.
CC InterPro; IPR004429; LeuB.
CC Pfam; PF00180; isodh; 1.
CC TIGRfam; TIGR00169; leuB; 1.
CC PROSITE; PS00470; IDH_IMDH; 1.
CC Oxidoreductase; Leucine biosynthesis; NAD; Complete proteome.
SQ SEQUENCE 370 AA; 39586 MW; 49566B394486D7FD CRC64;
Query Match 4.8%; Score 7; DB 1; Length 370;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 DLEKAI 61
| | | | |
Db 326 DLEKAI 332

RESULT 7
PRIL_ARCFU STANDARD; PRT; 370 AA.
ID PRIL_ARCFU
AC O29911;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

DE   Probable DNA primase large subunit (EC 2.7.7.-).
GN   PRIB OR AF0336.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1].
RS   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RC   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton P.M., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Gayland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus";
RL   Nature 390:364-370(1997).
CC   -!- FUNCTION: DNA primase is the polymerase that synthesizes small RNA
CC   primers for the Okazaki fragments on both template strands at
CC   replication forks during chromosomal DNA synthesis (By
CC   similarity).
CC   -!- SUBUNIT: Heterodimer of a small subunit and a large subunit (By
CC   similarity).
CC   -!- SIMILARITY: Belongs to the eukaryotic-type primase large subunit
CC   family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL; AE001081; AAB90898.1; -.
DR   DR
DR   PIR; H69291; H69291.
DR   TIGR; AF0336; -.
DR   HAMAP; MF 00701; -.
DR   InterPro; IPR007238; DNA_primase_lrg.
DR   Pfam; PF04104; DNA_primase_lrg; 1.
DR   Transfrase; DNA_replication; DNA-directed RNA polymerase; Primosome;
KW   Complete proteome.
SQ   SEQUENCE 370 AA; 42844 MW; 4369F68516EB90E CRC64;
Query Match 4.8%; Score 7; DB 1; Length 370;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY 104 REILSEL 110
Db |||||
Db 270 REILSEL 276
RESULT 8
YHG3_YEAST STANDARD; PRT; 429 AA.
AC F38756;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 48.9 kDa protein in RPL14B-GPAL intergenic region.
GN YHR003C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

```

```

RN  SEQUENCE FROM N.A.
RP  Kidd K.K., Busygina V., Demille M.M.C., Speed W.C., Ruggeri V.,
RA  Kidd J.R., Pakstis A.J.;
RT  "Overall linkage disequilibrium in 33 populations for highly
RL  informative multistate haplotypes spanning the HOXB gene cluster.";
RN  Am. J. Hum. Genet. 67:235-235(2000).
RN  [4]
RP  SEQUENCE OF 188-253 FROM N.A.
RC  TISSUE=Placenta;
RX  MEDLINE=89378558; PubMed=2570724;
RA  Glampao A., Acampora D., Zappavigna V., Pannese M.,
RA  D'Esposito M., Care A., Faiella A., Stornaiuolo A., Russo G.,
RA  Simeone A., Boncinelli E., Paschle C.;
RT  "Differential expression of human HOX-2 genes along the anterior-
RT  posterior axis in embryonic central nervous system.";
RL  Differentiation 40:191-197(1989).
RN  [5]
RP  SEQUENCE OF 188-253 FROM N.A.
RC  MEDLINE=90215256; PubMed=2576652;
RX  Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
RA  Gaudino G., Stornaiuolo A., Caffero M., Faiella A., Simeone A.;
RT  "Organization of human class I homeobox genes.";
RL  Genome 31:745-756(1989).
CC  -!- FUNCTION: Sequence-specific transcription factor which is part of
CC  a developmental regulatory system that provides cells with
CC  specific positional identities on the anterior-posterior axis.
CC  -!- SUBCELLULAR LOCATION: Nuclear.
CC  -!- DEVELOPMENTAL STAGE: Expressed in whole embryos and fetuses at 5-9
CC  weeks from conception.
CC  -!- SIMILARITY: Belongs to the Antp homeobox family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sb-sib.ch).
CC  -----
DR  EMBL; X16667; CAA34657.1; -
DR  EMBL; U9298; AAD10852.1; -
DR  EMBL; AF287967; AAG31555.1; -
DR  EMBL; X16175; CAA34297.1; -
DR  PIR; S07543; WJHU2G.
DR  HSSP; P02833; 1SAN.
DR  TRANSFAC; T01723; -
DR  Genew; HGNC:5114; HOXB3.
DR  MIM; 142966; -
DR  InterPro; IPR001827; Antennapedia.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; homeobox; 1.
DR  PRINTS; PR00025; ANTENNAPEDIA.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00032; ANTENNAPEDIA; 1.
DR  PROSITE; PS00071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW  Transcription regulation.
FT  SITE 129 134
FT  DOMAIN 154 178 ANTP-TYPE HEXAPEPTIDE.
FT  DOMAIN 188 247 GLY-RICH.
FT  DNA_BIND 199 200 HOMEBOX.
FT  CONFLICT 199 200 QL -> HV (IN REF. 2).
SQ  SEQUENCE 431 AA; 44344 MW; 941706EDCC2975E5 CRC64;

Query Match 4.8%; Score 7; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 PSAAPTS 27
|||||

```

Db 92 PSAAPTS 98

```

RESULT 10
PROA_SYNEL
ID PROA_SYNEL STANDARD; PRT; 439 AA.
AC Q8DKU1;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-
DE semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde
DE dehydrogenase) (GSA dehydrogenase).
GN PROA OR TLR0764.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2225144; PubMed=12240834;
RX Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
CC -!- FUNCTION: Catalyzes the NADPH dependent reduction of L-gamma-
CC glutamyl 5-phosphate into L-glutamate 5-semialdehyde and
CC phosphate. The product spontaneously undergoes cyclization to form
CC 1-pyrroline-5-carboxylate.
CC -!- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate +
CC NADPH(+) = L-gamma-glutamyl 5-phosphate + NADPH.
CC -!- PATHWAY: Proline biosynthesis; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the gamma-glutamyl phosphate reductase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; AF005371; BAC08315.1; -
DR HAMAP; MF_00412; -; 1.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000965; Gglut_pp_reduct.
DR Pfam; PF00171; aldedh; 1.
DR TIGRPFAM; TIGR00407; proA; 1.
DR PROSITE; PS01223; PROA; 1.
KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ SEQUENCE 439 AA; 46857 MW; 7036EBC5C62245EC CRC64;

Query Match 4.8%; Score 7; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 DLEKAIA 61
|||||
DR 244 DLEKAIA 250

```

```

RESULT 11
GATB_THETH
ID GATB_THETH STANDARD; PRT; 469 AA.
AC O9LCX2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B
DE (EC 6.3.5.-) (Asp/Glu-ADT subunit B).
GN GATB.
OC Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
OX [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=20374936; PubMed=10913601;
RA Becker H.D., Min B., Jacobi C., Raczniak G., Pelaschier J., Roy H.,
RA Klein S., Kern D., Soell D.;
RT "The heterotrimeric Thermus thermophilus Asp-tRNA(Asn)
RT amidotransferase can also generate Gln-tRNA(Gln).";
RL FEBS Lett. 476:140-144(2000).
CC -1- FUNCTION: allows the formation of correctly charged Asn-tRNA(Asn)
CC or Gln-tRNA(Gln) through the transamidation of misacylated Asp-
CC tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both
CC of asparaginyl-tRNA or glutamyl-tRNA synthetases. The reaction
CC takes place in the presence of glutamine and ATP through an
CC activated phospho-Asp-tRNA(Asn) or phospho-Glu-tRNA(Gln).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Asn) + L-glutamine = ADP
CC + phosphate + L-glutamyl-tRNA(Gln) + L-glutamate.
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartyl-tRNA(Asn) + L-glutamine = ADP
CC + phosphate + L-asparaginyl-tRNA(Asn) + L-glutamate.
CC -1- SUBUNIT: Heterotrimer of A, B and C subunits (By similarity).
CC -1- SIMILARITY: Belongs to the gatB/gatE family. GatB subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF202448; AAF91177.1; -.
DR HAMAP; MF_00121; -.
DR InterPro; IPR004413; GatB.
DR InterPro; IPR006107; GatB_cent.
DR InterPro; IPR006075; GatB_N.
DR InterPro; IPR003789; GatB_Yqey.
DR Pfam; PF01162; GatB_1.
DR Pfam; PF02934; GatB_N; 1.
DR Pfam; PF02637; GatB_Yqey; 1.
DR TIGRFAMs; TIGR00133; gatB; 1.
DR PROSITE; PS01234; GATB; 1.
DR Protein biosynthesis; Ligase.
KW PROTEIN BIOSYNTHESIS; Ligase.
SQ SEQUENCE 469 AA; 52236 MW; 10D01DFF9C9B0625 CRC64;
Query Match 4.8%; Score 7; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 KSLHLEG 17
DB 125 KSLHLEG 131
RESULT 12
C71G ARATH
ID C71G ARATH STANDARD; PRT; 497 AA.
AC Q9PH66;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 71A16 (EC 1.14.-.-).
DE CYP71A16 OR AT5G42590 OR K16E1.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=5274;
OX [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=20374936; PubMed=10913601;
RA Becker H.D., Min B., Jacobi C., Raczniak G., Pelaschier J., Roy H.,
RA Klein S., Kern D., Soell D.;
RT "The heterotrimeric Thermus thermophilus Asp-tRNA(Asn)
RT amidotransferase can also generate Gln-tRNA(Gln).";
RL FEBS Lett. 476:140-144(2000).
CC -1- FUNCTION: allows the formation of correctly charged Asn-tRNA(Asn)
CC or Gln-tRNA(Gln) through the transamidation of misacylated Asp-
CC tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both
CC of asparaginyl-tRNA or glutamyl-tRNA synthetases. The reaction
CC takes place in the presence of glutamine and ATP through an
CC activated phospho-Asp-tRNA(Asn) or phospho-Glu-tRNA(Gln).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Asn) + L-glutamine = ADP
CC + phosphate + L-glutamyl-tRNA(Gln) + L-glutamate.
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartyl-tRNA(Asn) + L-glutamine = ADP
CC + phosphate + L-asparaginyl-tRNA(Asn) + L-glutamate.
CC -1- SUBUNIT: Heterotrimer of A, B and C subunits (By similarity).
CC -1- SIMILARITY: Belongs to the gatB/gatE family. GatB subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF202448; AAF91177.1; -.
DR HAMAP; MF_00121; -.
DR InterPro; IPR004413; GatB.
DR InterPro; IPR006107; GatB_cent.
DR InterPro; IPR006075; GatB_N.
DR InterPro; IPR003789; GatB_Yqey.
DR Pfam; PF01162; GatB_1.
DR Pfam; PF02934; GatB_N; 1.
DR Pfam; PF02637; GatB_Yqey; 1.
DR TIGRFAMs; TIGR00133; gatB; 1.
DR PROSITE; PS01234; GATB; 1.
DR Protein biosynthesis; Ligase.
KW PROTEIN BIOSYNTHESIS; Ligase.
SQ SEQUENCE 469 AA; 52236 MW; 10D01DFF9C9B0625 CRC64;
Query Match 4.8%; Score 7; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 KSLHLEG 17
DB 125 KSLHLEG 131
RESULT 13
SH2B MOUSE
ID SH2B MOUSE STANDARD; PRT; 504 AA.
AC Q02152;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-hydroxytryptamine 2B receptor (5-HT-2B) (serotonin receptor)
DE (5-HT-2F) (NP75 protein).
DE HTR2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93050195; PubMed=1426253;
RA Loric S., Launay J.-M., Colas J.-F., Maroteaux L.;
RT "New mouse 5-HT2-like receptor. Expression in brain, heart and
RT intestine.";
RL FEBS Lett. 312:203-207(1992).
CC -1- FUNCTION: This is one of the several different receptors for 5-
CC hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. This receptor
CC mediates its action by association with G proteins that activate a
CC phosphatidylinositol-calcium second messenger system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Intestine and heart, but also in brain and
CC kidney.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Strongest to the other 5HT-2 subtype receptors.

```



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z15119; CAA78824.1; -
CC FR: S27269; S27269.
CC MG: MGI109323; Rtr2b.
CC GO: GO:0007507; P:heart development; IMP.
CC InterPro: IPR000276; GPCR_Rhodopsin.
CC Pfam: PF00001; 7tm 1; 1.
CC PRINTS: PR00237; GPCRHDOPSN.
CC PROSITE: PS00237; G PROTEIN RECEPT_F1_1; 1.
CC PROSITE: PS00262; G PROTEIN RECEPT_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 55 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 56 78 1 (POTENTIAL).
FT DOMAIN 79 90 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 91 112 2 (POTENTIAL).
FT DOMAIN 113 128 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 129 150 3 (POTENTIAL).
FT DOMAIN 151 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 191 4 (POTENTIAL).
FT DOMAIN 192 215 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 216 238 5 (POTENTIAL).
FT DOMAIN 239 323 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 324 344 6 (POTENTIAL).
FT DOMAIN 345 359 7 (POTENTIAL).
FT TRANSMEM 360 382 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 383 504 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 127 206 BY SIMILARITY.
FT LIPID 396 396 S-palmitoyl cysteine (Potential).
SQ SEQUENCE 504 AA; 56545 MW; 045A483E281ECD30 CRC64;

Query Match 4.8%; Score 7; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 SVKALRK 52
DB 403 SVKALRK 409

RESULT 14
ID SYH_FUGRU STANDARD; PRT; 519 AA.
AC P70076;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE (HISRS).
GN HARS OR HISS.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96323249; PubMed=8710896;
RA Brenner S., Corrochano L.M.;
RT "translocation events in the evolution of aminoacyl-tRNA
RT synthetases."
RL Proc. Natl. Acad. Sci. U.S.A. 93:8485-8489(1996).

```

```

CC -----
CC -!- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC diphosphate + L-histidyl-tRNA(His).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z54243; CAA91012.1; -
CC InterPro: IPR004154; HGTP_anticonodon.
CC InterPro: IPR004516; Hiss.
CC InterPro: IPR002314; tRNA-synt_2b.
CC InterPro: IPR006195; tRNA_ligase_II.
CC InterPro: IPR00738; WHEP-TRS.
CC Pfam: PF03129; HGTP_anticonodon; 1.
CC Pfam: PF00587; tRNA-synt_2b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC TIGRFAMs: TIGR00442; Hiss; 1.
CC PROSITE: PS50862; AA tRNA_LIGASE_II; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
KW Aminoacyl-tRNA synthetase; 57913 MW; AICBF5752070759E CRC64;
SQ SEQUENCE 519 AA; 57913 MW; AICBF5752070759E CRC64;

Query Match 4.8%; Score 7; DB 1; Length 519;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 EILSELD 111
DB 202 EILSELD 208

RESULT 15
ID DPAL_SUNY3 STANDARD; PRT; 573 AA.
AC Q55393;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Diflavin flavoprotein A 1 (EC 1.-.-.) (SBATF573) (NADH:oxygen
DE oxidoreductase).
DE DPAL OR SLL0550.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RX Kaneko T., Tanaka A., Sato S., Kotani H., Suzuki T., Miyajima N.,
RA Sugiyama M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome."
RL DNA Res. 2:153-166(1995).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98321624; PubMed=9660187;
RX Wasserfallen A., Ragettli S., Jouanneau Y., Leisinger T.;
RT "A family of flavoproteins in the domains Archaea and Bacteria."
RL Eur. J. Biochem. 254:325-332(1998).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=22050187; PubMed=12054744;
RX Vicente J.B., Gomes C.M., Wasserfallen A., Teixeira M.;
RT "Module fusion in an A-type flavoprotein from the cyanobacterium
RT Synecocystis condenses a multiple-component pathway in a single
RT polypeptide chain."
RL Biochem. Biophys. Res. Commun. 294:82-87(2002).
CC -!- FUNCTION: Mediates electron transfer from NADH to oxygen, reducing

```

CC it to water. This modular protein has 3 redox cofactors, in other
CC organisms the same activity requires 2 or 3 proteins.
CC -|- COPACTOR: Binds 2 iron atoms and a variable amount of FAD and FMN
CC per monomer.
CC -|- SUBUNIT: Homodimer.
CC -|- MISCELLANEOUS: By homology with norV in E.coli, may be involved in
CC nitric oxide detoxification (By similarity).
CC -|- SIMILARITY: In the N-terminal section; belongs to the zinc
CC metallo-hydrolase family group 3.
CC -|- SIMILARITY: In the C-terminal section; belongs to the flavodoxin
CC reductase family.
CC -|- SIMILARITY: Contains 1 flavodoxin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D64003; BAA10483.1; -.
CC PIR; S75748; S75748.
CC InterPro; IPR001279; Blactmase-like.
CC InterPro; IPR008254; Flav_nitox_synth.
CC InterPro; IPR002583; Flav_nitox_synth.
CC InterPro; IPR001226; Flavodoxin.
CC InterPro; IPR009002; FMN binding.
CC Pfam; PF01613; Flav_nitox_synth.
CC Pfam; PF0753; Lactamase B; 1.
CC PROSITE; PS50902; FLAVODOXIN LIKE; 1.
KW Transport; Electron transport; Oxidoreductase; Flavoprotein; FAD; FMN;
KW Metal-binding; Iron; Multigene family; Complete proteome.
FT DOMAIN 43 236 ZINC METALLO-HYDROLASE.
FT DOMAIN 265 401 FLAVODOXIN-LIKE.
FT DOMAIN 424 573 FLAVODOXIN-REDUCTASE-LIKE.
FT METAL 92 92 IRON 1 (BY SIMILARITY).
FT METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 96 96 IRON 2 (BY SIMILARITY).
FT METAL 159 159 IRON 1 (BY SIMILARITY).
FT METAL 178 178 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 236 236 IRON 2 (BY SIMILARITY).
SQ SEQUENCE 573 AA; 63506 MW; E31BE51028949E58 CRC64;

Query Match 4.8%; Score 7; DB 1; Length 573;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 SDLEKAI 60
Db 418 SDLEKAI 424

Search completed: June 3, 2004, 13:29:30
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 13:30:28 ; Search time 43 Seconds
(without alignments)
961.786 Million cell updates/sec

Title: US-09-744-197-1
Perfect score: 147
Sequence: 1 MGGCMHSTQDKSLHLEGPDPN.....LSITVMSDLLQNIIRNVKIMK 147

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1155919 seqs, 28135677 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	93.9	140	9	US-09-764-864-1447
2	138	5.4	314	9	US-09-764-864-1447
3	7	4.8	41	9	US-09-764-864-1447
4	7	4.8	123	12	US-10-425-114-51057
5	7	4.8	127	10	US-09-764-864-1447
6	7	4.8	162	15	US-10-425-114-51057
7	7	4.8	173	12	US-10-425-114-51057
8	7	4.8	180	12	US-10-425-114-51057
9	7	4.8	183	15	US-10-425-114-51057
10	7	4.8	188	14	US-10-425-114-51057
11	7	4.8	213	15	US-10-425-114-51057
12	7	4.8	245	16	US-10-425-114-51057
13	7	4.8	378	15	US-10-425-114-51057
14	7	4.8	414	12	US-10-425-114-51057
15	7	4.8	417	12	US-10-425-114-51057

16	7	4.8	496	12	US-10-425-114-51057
17	7	4.8	504	10	US-09-903-376-2
18	7	4.8	552	12	US-10-425-114-58950
19	7	4.8	747	12	US-10-072-012-566
20	7	4.8	790	15	US-10-369-493-21510
21	7	4.8	804	16	US-10-389-566-1940
22	7	4.8	890	15	US-10-369-493-23121
23	7	4.8	1311	8	US-08-954-701A-4
24	7	4.8	1311	10	US-09-754-032-4
25	7	4.8	1311	14	US-10-421-446-4
26	7	4.8	19723	15	US-10-084-846A-5
27	6	4.1	9	10	US-09-793-451-48
28	6	4.1	9	10	US-09-793-451-263
29	6	4.1	9	10	US-09-793-451-364
30	6	4.1	9	10	US-09-793-451-415
31	6	4.1	9	14	US-10-283-722-48
32	6	4.1	9	14	US-10-283-722-263
33	6	4.1	9	14	US-10-283-722-364
34	6	4.1	9	14	US-10-283-722-415
35	6	4.1	9	15	US-10-283-903-48
36	6	4.1	9	15	US-10-283-903-263
37	6	4.1	9	15	US-10-283-903-364
38	6	4.1	9	15	US-10-283-903-415
39	6	4.1	9	15	US-10-428-335-50
40	6	4.1	9	15	US-10-428-335-135
41	6	4.1	10	10	US-09-793-451-398
42	6	4.1	10	10	US-09-793-451-499
43	6	4.1	10	10	US-09-793-451-503
44	6	4.1	10	10	US-09-793-451-605
45	6	4.1	10	10	US-09-793-451-693

ALIGNMENTS

RESULT 1
US-09-764-864-1447
; Sequence 1447, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1447
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1447

Query Match 93.9%; Score 138; DB 9; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.9e-127;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 DKSLHLEGPDPNPSAAPTSTCAPRKPVKRISISKQLASVKALRKCSDLEKATATTALIPRN 69
DB 3 DKSLHLEGPDPNPSAAPTSTCAPRKPVKRISISKQLASVKALRKCSDLEKATATTALIPRN 62
QY 70 SSDSGKLEKAIKADLLQTFRNFAEGQETPKYREILSELDEHTENKLDFFEDFMILL 129
DB 63 SSDSGKLEKAIKADLLQTFRNFAEGQETPKYREILSELDEHTENKLDFFEDFMILL 122
QY 130 ITVMSDLLQNIIRNVKIMK 147
DB 123 ITVMSDLLQNIIRNVKIMK 140

RESULT 2
US-09-978-274A-2

```

, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/508,408
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/774,203
, PRIOR FILING DATE: 2001-01-29
, NUMBER OF SEQ ID NOS: 49117
, SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
, SEQ ID NO 41438
, LENGTH: 41
, TYPE: PR1
, ORGANISM: Homo sapiens
, FEATURE:
, OTHER INFORMATION: MAP TO AC015525.3
, OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3
, OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
, OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
, OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
, OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.7
, OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
, OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
, OTHER INFORMATION: EST_HUMAN HIT: BE789681.1, EVALUATE 1.00e-07
, US-09-864-761-41438
Query Match 4.8%; Score 7; DE 9; Length 41;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 2; Conservativeness 0; Mismatches 0; Indels

```

```

RESULT 4
US-10-424-599-199526
; Sequence 199526, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199526
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(123)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22197C.1.pep
; US-10-424-599-199526

Query Match 4.8%; Score 7; DB 12; Length 123;
Best Local Similarity 100.0%; Pred.No. 87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 ALRKCS D 55
Db 106 ALRKCS D 112

RESULT 5
US-09-764-891-4168
; Sequence 4168, Application US/09764891

```

Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4168
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-891-4168

Query Match 4.8%; Score 7; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 SAAPTST 28
Db 90 SAAPTST 96

RESULT 6
US-10-094-749-2625
Sequence 2625, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2625
LENGTH: 162
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2625

Query Match 4.8%; Score 7; DB 15; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 LASVKAL 50
Db 92 LASVKAL 98

RESULT 7
US-10-425-114-39200
Sequence 39200, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 39200
LENGTH: 173
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3079-017-B11_FLI.pep
US-10-425-114-39200

Query Match 4.8%; Score 7; DB 12; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 GKLEKAI 81
Db 76 GKLEKAI 82

RESULT 8
US-10-425-114-55789
Sequence 55789, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 55789
LENGTH: 180
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4698-005-G5_FLI.pep
US-10-425-114-55789

Query Match 4.8%; Score 7; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 GKLEKAI 81
Db 79 GKLEKAI 85

RESULT 9
US-10-369-493-22366
Sequence 22366, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
PRIOR FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 22366
LENGTH: 183
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22366

Query Match 4.8%; Score 7; DB 15; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 ATTALIF 67
Db 84 ATTALIF 90

RESULT 10
US-10-106-698-5287
Sequence 5287, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005PI
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 5287
LENGTH: 188
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: MISC FEATURE
LOCATION: (170)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (172)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (177)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (183)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (188)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5287

Query Match 4.8%; Score 7; DB 14; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ILSDELDE 112

Db 46 ILSDELDE 52

RESULT 11
US-10-393-840-53
Sequence 53, Application US/10393840
Publication No. US20030229922A1
GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
TITLE OF INVENTION: Materials and Methods for the
TITLE OF INVENTION: Modification of Plant Cell Wall Polysaccharides
FILE REFERENCE: 11000.1012c3
CURRENT APPLICATION NUMBER: US/10/393,840
CURRENT FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: US 09/636,800
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 09/170,862
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: US 60/148,426
PRIOR FILING DATE: 1999-08-11
PRIOR APPLICATION NUMBER: PCT NZ/99/00169
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 956
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 53
LENGTH: 213
TYPE: PRT
ORGANISM: Pinus radiata
US-10-393-840-53

Query Match 4.8%; Score 7; DB 15; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GKLEKAI 81
Db 73 GKLEKAI 79

RESULT 12
US-10-389-566-398
Sequence 398, Application US/10389566
Publication No. US20040025202A1
GENERAL INFORMATION:
APPLICANT: Monsanto Technology, LLC
APPLICANT: Laurie, Cathy C
TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
FILE REFERENCE: 38-77(52900)D
CURRENT APPLICATION NUMBER: US/10/389,566
CURRENT FILING DATE: 2003-03-31
PRIOR APPLICATION NUMBER: US 60/365,301
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/391,786
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/392,018
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 2459
SOFTWARE: PatentIn version 3.2
SEQ ID NO 398
LENGTH: 245
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
LOCATION: (21)..
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-398

Query Match 4.8%; Score 7; DB 16; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 GKLEKAI 81
|||||
Db 91 GKLEKAI 97

RESULT 13

US-10-369-493-17232
; Sequence 17232, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17232
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-10-369-493-17232

Query Match 4.8%; Score 7; DB 15; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 LEKAIAT 62
|||||
Db 58 LEKAIAT 64

RESULT 14

US-10-425-114-52749
; Sequence 52749, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52749
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700562414_FLI.pep
US-10-425-114-52749

Query Match 4.8%; Score 7; DB 12; Length 414;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 DGKLEKA 80
|||||
Db 350 DGKLEKA 356

RESULT 15

US-10-424-599-167317
; Sequence 167317, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167317
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_122104C.1.pep
US-10-424-599-167317

Query Match 4.8%; Score 7; DB 12; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 DGKLEKA 80
|||||
Db 353 DGKLEKA 359

Search completed: June 3, 2004, 13:36:14
Job time : 43 secs